

STIC-Biotech/ChemLib

From: Mertz, Prema
Sent: Thursday, March 02, 2006 2:44 PM
To: STIC-Biotech/ChemLib
Subject: 10/695,994

Please search any 6 amino acids of SEQ ID NO:7 with protein databases.

Thanks.

Prema Mertz, Ph.D., J.D.
Primary Examiner
Art Unit 1646
4D81 Remsen Bldg Mailbox 4C70
US Patent & Trademark Office
Tel # (571) 272-0876
FAX # (571) 273-0876

18/105 me
RECEIVED
MAR - 2 2006
STIC

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2006, 07:02:06 ; Search time 187 Seconds
(without alignments)
516.916 Million cell updates/sec

Title: US-10-695-994A-7

Perfect score: 220

Sequence: 1 MEIQRITSSISGPIISPAYTQ.....QNYKFNRSRRTPADHSWSGM 220

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2443163 seqs, 439378781 residues

Word size: 6

Total number of hits satisfying chosen parameters: 4830

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

1: Geneseqp21:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	220	100.0	220	7	ADP74494
2	220	100.0	220	8	ADQ88366
3	129	58.6	220	2	AAW81640
4	124	56.4	1022	8	ADP13682
5	124	56.4	2154	2	AAW81639
6	124	56.4	2154	7	ADP74491
7	124	56.4	2154	9	ADQ88363
8	124	56.4	2154	9	ADZ85101
9	97	44.1	321	6	ABU70467
10	97	44.1	1630	7	ADJ70625
11	97	44.1	2364	5	ABP5068
12	97	44.1	2364	8	ABM80299
13	97	44.1	2364	8	ADZ70337
14	97	44.1	2364	9	ADZ70337
15	90	40.9	1120	2	AAW81642
16	38	17.3	2106	7	ADJ70287
17	38	17.3	2137	9	ADX06608
18	38	17.3	2141	6	ABR41636
19	38	17.3	2387	4	AAU01183
20	27	12.3	2390	2	AAV05494
21	23	10.5	806	4	ABP41709
22	19	8.6	2291	4	ABB61876
23	15	6.8	15	7	ADP74518
24	15	6.8	15	8	ADQ88396

25	15	6.8	17	9	ADV13561	ADV13561 Human pho
26	15	6.8	17	9	ADV13562	ADV13562 Human pho
27	14	6.4	14	2	AAW81644	AAW81644 Mouse elf
28	14	6.4	14	7	ADP74507	ADP74507 Peptide f
29	14	6.4	14	8	ADQ88379	ADQ88379 Mouse elf
30	14	6.4	14	8	ADY65644	ADY65644 S. manson
31	14	6.4	2326	4	AAW50652	AAW50652 C. elegans
32	14	6.4	2326	8	ADN23721	ADN23721 Bacterial
33	11	5.0	385	9	AAW5025	AAW5025 Human act
34	11	5.0	897	3	AAW59242	AAW59242 A rod sho
35	11	5.0	900	3	AAW59241	AAW59241 A rod sho
36	11	5.0	1092	3	AAW59237	AAW59237 A rod sho
37	11	5.0	1201	3	AAW59238	AAW59238 A rod sho
38	11	5.0	1310	3	AAW59239	AAW59239 A rod sho
39	11	5.0	1310	3	AAW59239	AAW59239 A rod sho
40	11	5.0	1355	8	ADN22598	ADN22598 Bacterial
41	11	5.0	1355	8	ADN22599	ADN22599 Bacterial
42	11	5.0	2000	9	AEA24046	AEA24046 Human PRO
43	11	5.0	3685	1	AAW90373	AAW90373 Sequence.
44	11	5.0	3685	1	AAW90290	AAW90290 Human Duc
45	11	5.0	3685	8	ADL83110	ADL83110 Human PRO
46	11	5.0	3685	8	ADN04004	ADN04004 Antiprot
47	11	5.0	5373	4	AAU14603	AAU14603 Novel bon
48	11	5.0	5373	7	ADJ68935	ADJ68935 Human hea
49	11	5.0	5430	9	ADZ85102	ADZ85102 Full-leng
50	11	5.0	5447	4	AAU14697	AAU14697 Novel bon
51	11	5.0	7201	4	ABW71136	ABW71136 Drosophill
52	10	4.5	5385	4	ABW66487	ABW66487 Drosophill
53	10	4.5	5496	4	ABW67161	ABW67161 Drosophill
54	9	4.1	67	5	ABP62976	ABP62976 Peptide u
55	9	4.1	72	2	AAW93379	AAW93379 Canine BP
56	9	4.1	108	3	AAW01083	AAW01083 Human sec
57	9	4.1	117	4	ABG11927	ABG11927 Novel hum
58	9	4.1	201	1	ABP98499	ABP98499 Partial s
59	9	4.1	202	8	ADY04237	ADY04237 Plant ful
60	9	4.1	205	8	ADY09483	ADY09483 Plant ful
61	9	4.1	240	8	ADX66116	ADX66116 Plant ful
62	9	4.1	474	4	ABW67963	ABW67963 Amino aci
63	9	4.1	2000	8	ADP55039	ADP55039 Human PRO
64	9	4.1	2008	2	AAW22016	AAW22016 Utrrophin
65	9	4.1	2013	4	AAW67964	AAW67964 Amino aci
66	9	4.1	2105	8	ADQ44170	ADQ44170 Structura
67	9	4.1	3064	6	ABO14706	ABO14706 Novel hum
68	9	4.1	3215	8	ADL11060	ADL11060 Human the
69	9	4.1	3227	8	ADL11061	ADL11061 Human the
70	9	4.1	3432	9	ADW10726	ADW10726 Dog utrop
71	9	4.1	3433	2	AAW22017	AAW22017 Utrrophin.
72	9	4.1	3433	7	ADJ68925	ADJ68925 Human hea
73	9	4.1	3433	7	ADJ68926	ADJ68926 Human hea
74	9	4.1	3433	8	ADP55624	ADP55624 Human PRO
75	9	4.1	4576	8	ADQ55184	ADQ55184 Protein #
76	9	3.6	358	6	ABJ25461	ABJ25461 Aspergill
77	8	3.6	517	7	ABM88463	ABM88463 Rice abio
78	8	3.6	517	6	ABJ26061	ABJ26061 Aspergill
79	8	3.6	547	4	ABW63123	ABW63123 Drosophill
80	8	3.6	592	8	ADN72217	ADN72217 Thale cre
81	8	3.6	626	6	ABR41653	ABR41653 Human DIT
82	8	3.6	690	4	ABW63142	ABW63142 Drosophill
83	8	3.6	884	5	ABG34125	ABG34125 Human alp
84	8	3.6	894	8	ADL24315	ADL24315 AW755252-
85	8	3.6	894	8	ADX26296	ADX26296 Novel cel
86	8	3.6	894	9	ADX26371	ADX26371 Novel cel
87	8	3.6	894	9	ADX26304	ADX26304 Rat Prote
88	8	3.6	911	7	ADW62302	ADW62302 Rat Prote
89	8	3.6	911	8	ADQ05018	ADQ05018 Pancreas
90	8	3.6	912	4	AAW90771	AAW90771 Human she
91	8	3.6	912	5	ABG34124	ABG34124 Human non
92	8	3.6	912	6	ABU89754	ABU89754 Protein d
93	8	3.6	912	6	ADL24316	ADL24316 AW755252-
94	8	3.6	937	4	AAW63240	AAW63240 Human bre
95	8	3.6	1190	8	ADW29513	ADW29513 Bacterial
96	8	3.6				
97	8	3.6				

98	8	3.6	1519	4	ABB66635	Abb66635 Drosophill	171	7	3.2	267	8	ADL91559	ADL91559 Human tmn
99	8	3.6	2129	4	ABG15478	Abg15478 Novel hum	172	7	3.2	271	8	ADM32828	Adm32828 Amrno aci
100	8	3.6	2129	4	ABG20749	Abg20749 Novel hum	173	7	3.2	274	8	AAAG06912	AAAG06912 Arabidops
101	8	3.6	2414	2	AA84882	Aa84882 Transcrip	174	7	3.2	274	3	AAAG53993	AAAG53993 Arabidops
102	8	3.6	2414	2	AAW4057	Aaw4057 Cellular	175	7	3.2	280	8	ADX89926	Adx89926 Plant ful
103	8	3.6	2414	5	ABW06340	Abw06340 Human p30	176	7	3.2	290	8	ADT58314	Adt58314 Plant pol
104	8	3.6	2414	6	ABU03968	Abu03968 Human exp	177	7	3.2	291	2	AA868128	Aa868128 Human Mac
105	8	3.6	2414	6	ABU03980	Abu03980 Human exp	178	7	3.2	295	7	ADC39152	Adc39152 Novel hum
106	8	3.6	2414	6	ABU03975	Abu03975 Human exp	179	7	3.2	298	8	ADX78880	Adx78880 Plant ful
107	8	3.6	2414	6	ABU03978	Abu03978 Human exp	180	7	3.2	300	8	ADU01159	Adu01159 Human pro
108	8	3.6	2414	6	ABU03976	Abu03976 Human exp	181	7	3.2	300	8	ADU15541	Adu15541 Novel hum
109	8	3.6	4097	6	ABW57815	Abw57815 Drosophill	182	7	3.2	302	9	ADM17390	Adm17390 Eucalyptu
110	8	3.6	4101	8	ADN23857	Adn23857 Bacterial	183	7	3.2	306	4	ABG03251	Abg03251 Novel hum
111	8	3.6	4101	8	ADN23856	Adn23856 Bacterial	184	7	3.2	306	5	ABP38671	Abp38671 Stephyloc
112	7	3.2	9	4	AAE01104	Aae01104 Human leu	185	7	3.2	306	8	ADSO5186	Adso5186 Stephyloc
113	7	3.2	13	2	AAW81643	Aaw81643 Mouse elf	186	7	3.2	307	8	ADU02746	Adu02746 Novel hum
114	7	3.2	13	7	ADP74506	Adp74506 Peptide f	187	7	3.2	307	8	ADQ66295	Adq66295 Novel hum
115	7	3.2	13	8	ADQ88378	Adq88378 Mouse elf	188	7	3.2	309	8	ADQ66295	Adq66295 Novel hum
116	7	3.2	13	8	ADQ88395	Adq88395 Mouse elf	189	7	3.2	314	3	AAV91560	Aav91560 Human sec
117	7	3.2	32	7	ADDO6181	Ado6181 Yeast MAB	190	7	3.2	314	3	AAV91672	Aav91672 Human sec
118	7	3.2	55	8	ABO56432	AbO56432 Human gen	191	7	3.2	314	8	ADL71751	Adl71751 Novel hum
119	7	3.2	67	4	AAU46339	Aau46339 Propionib	192	7	3.2	316	8	ADL71635	Adl71635 Novel hum
120	7	3.2	67	6	ABW42858	Abw42858 Propionib	193	7	3.2	316	5	ABP29264	Abp29264 Streptoco
121	7	3.2	74	2	AAV33806	Aav33806 YUL124C a	194	7	3.2	316	5	ABP29264	Abp29264 Streptoco
122	7	3.2	91	7	ADU70926	Adu70926 Human hea	195	7	3.2	321	6	ABR43288	AbR43288 Human neu
123	7	3.2	95	4	AAU48469	Aau48469 Propionib	196	7	3.2	326	2	AAV20109	Aav20109 B. burgdo
124	7	3.2	95	6	ABW44988	Abw44988 Propionib	197	7	3.2	326	4	ABW58880	Abw58880 Drosophill
125	7	3.2	100	6	ABW66150	Abw66150 Novel hum	198	7	3.2	344	4	ABW58880	Abw58880 Drosophill
126	7	3.2	109	4	ABG26491	Abg26491 Novel hum	199	7	3.2	344	7	ADC39150	Adc39150 Novel hum
127	7	3.2	115	3	ABW58937	Abw58937 Breast an	200	7	3.2	346	7	AAV20108	Aav20108 B. burgdo
128	7	3.2	120	4	AAW18063	Aaw18063 Peptide #	201	7	3.2	347	2	ADC39146	Adc39146 Novel hum
129	7	3.2	120	4	ABW37101	Abw37101 Peptide #	202	7	3.2	352	7	ADQ66129	Adq66129 Novel hum
130	7	3.2	120	4	AAW30575	Aaw30575 Peptide #	203	7	3.2	357	3	AAAG53992	AAAG53992 Arabidops
131	7	3.2	120	4	ABW31865	Abw31865 Peptide #	204	7	3.2	358	8	AAAG53992	AAAG53992 Arabidops
132	7	3.2	120	4	ABW22409	Abw22409 Protein #	205	7	3.2	358	5	ABW92618	Abw92618 Herbicida
133	7	3.2	120	4	AAW70237	Aaw70237 Human bra	206	7	3.2	358	5	AAE31707	Aae31707 Flavonone
134	7	3.2	120	4	AAW57823	Aaw57823 Human Dna	207	7	3.2	358	6	AAE31707	Aae31707 Flavonone
135	7	3.2	120	4	ABW51934	Abw51934 Human liv	208	7	3.2	358	9	ADZ25324	Adz25324 Human Mac
136	7	3.2	120	4	AAW5702	Aaw5702 Peptide #	209	7	3.2	360	2	AAV33492	Aav33492 Human neu
137	7	3.2	120	5	ABG39875	Abg39875 Human pep	210	7	3.2	362	6	ABR43289	AbR43289 Human neu
138	7	3.2	132	4	ABW11597	Abw11597 Human sec	211	7	3.2	365	6	ABG74623	Abg74623 Parsley F
139	7	3.2	132	6	ABW11665	Abw11665 Human MMD	212	7	3.2	366	6	ABU94255	Abu94255 Trifolium
140	7	3.2	140	8	ADW58759	Adw58759 Protein f	213	7	3.2	368	7	ADC39148	Adc39148 Novel hum
141	7	3.2	156	5	ABP51297	Abp51297 Human MMD	214	7	3.2	368	7	ADQ65520	Adq65520 Human ata
142	7	3.2	158	4	ABW53284	Abw53284 Human tes	215	7	3.2	373	3	ABU70528	Abu70528 Human adi
143	7	3.2	158	6	ADW10654	Adw10654 Allolococ	216	7	3.2	384	6	ABU70528	Abu70528 Human adi
144	7	3.2	158	7	ADW10623	Adw10623 Secreted	217	7	3.2	384	7	ABO75340	AbO75340 Pseudomon
145	7	3.2	160	7	ABO79500	AbO79500 Pseudomon	218	7	3.2	385	7	ADP59634	Adp59634 Human pol
146	7	3.2	160	3	AAW38862	Aaw38862 Arabidops	219	7	3.2	392	3	AAW29453	Aaw29453 Arabidops
147	7	3.2	168	8	ADU02925	Adu02925 Novel hum	220	7	3.2	397	5	ABW48741	Abw48741 Bacteria
148	7	3.2	172	6	ABW53405	Abw53405 Protein s	221	7	3.2	399	8	ADW24890	Adw24890 Bacteria
149	7	3.2	172	6	ADW10652	Adw10652 Allolococ	222	7	3.2	401	7	ABO78163	AbO78163 Pseudomon
150	7	3.2	172	6	ADW64720	Adw64720 Disease t	223	7	3.2	412	3	AAW29452	Aaw29452 Arabidops
151	7	3.2	173	8	ADW67034	Adw67034 Novel hum	224	7	3.2	422	6	ABJ25881	Abj25881 Aspergill
152	7	3.2	186	8	ADW65572	Adw65572 Human pro	225	7	3.2	426	6	ABJ25881	Abj25881 Aspergill
153	7	3.2	186	8	ADW58909	Adw58909 Human Elk	226	7	3.2	426	6	ADW11404	Adw11404 Allolococ
154	7	3.2	199	7	ADW05425	Adw05425 Human pro	227	7	3.2	433	6	AAO30802	Aao30802 Human SH2
155	7	3.2	205	8	ADL14070	Adl14070 Human bar	228	7	3.2	438	7	ABO68604	AbO68604 Pseudomon
156	7	3.2	211	4	AAE01083	Aae01083 Human col	229	7	3.2	450	8	ADN60447	Adn60447 B. lichen
157	7	3.2	211	7	ABW88336	Abw88336 Rice abio	230	7	3.2	450	8	ADN60447	Adn60447 B. lichen
158	7	3.2	212	4	AAW63593	Aaw63593 Human gas	231	7	3.2	451	7	ABO78856	AbO78856 Pseudomon
159	7	3.2	218	5	ABG91582	Abg91582 Purine/py	232	7	3.2	453	6	ABW25446	Abw25446 Protein e
160	7	3.2	218	8	ABO58433	AbO58433 Human gen	233	7	3.2	454	4	ABG22866	Abg22866 Novel hum
161	7	3.2	218	8	ADW29497	Adw29497 Bacterial	234	7	3.2	456	4	AAW40115	Aaw40115 Human pol
162	7	3.2	221	4	AAW63595	Aaw63595 Human gas	235	7	3.2	465	4	AAW41901	Aaw41901 Human pol
163	7	3.2	229	8	ADW15444	Adw15444 Rhodococ	236	7	3.2	470	8	ADJ49948	Adj49948 Oll--aseoc
164	7	3.2	246	8	ADW2785	Adw2785 Plant ful	237	7	3.2	471	8	ADQ20174	Adq20174 Human pro
165	7	3.2	253	8	ADW21602	Adw21602 Bacterial	238	7	3.2	471	8	ADP25221	Adp25221 PRO polyp
166	7	3.2	254	3	AAW63594	Aaw63594 Arabidops	239	7	3.2	471	9	ADY17776	Ady17776 PRO polyp
167	7	3.2	254	3	AAW63593	Aaw63593 Arabidops	240	7	3.2	491	1	AAV70464	Aav70464 Sequence
168	7	3.2	258	3	AAW29454	Aaw29454 Arabidops	241	7	3.2	498	9	ADW39495	Adw39495 HIV Gag p
169	7	3.2	259	3	AAW56375	Aaw56375 Human pro	242	7	3.2	500	5	ABP73751	Abp73751 Candida a
170	7	3.2	262	7	ADW60612	Adw60612 Human con	243	7	3.2	505	7	ADD48723	Add48723 Human pro

244	7	3.2	524	5	ABP51657	Abp51657 Drosophila	317	7	3.2	956	7	ADM04352	Adm04352 Human pro
245	7	3.2	534	4	AA939350	AA939350 Human pro	318	7	3.2	957	4	AAW79016	AAW79016 Human pro
246	7	3.2	539	7	AD08421	Novel pro	319	7	3.2	959	4	AAW41353	AAW41353 Human pol
247	7	3.2	540	4	AA974992	AA974992 Human pol	320	7	3.2	1002	4	AAW39567	AAW39567 Human pol
248	7	3.2	548	8	ADN17887	ADN17887 Bacterial	321	7	3.2	1002	9	ABE56484	ABE56484 Radiocchem
249	7	3.2	549	7	AD65325	AD65325 Human pro	322	7	3.2	1045	4	ABE62186	ABE62186 Drosophila
250	7	3.2	555	7	ADP55438	ADP55438 Human nov	323	7	3.2	1045	6	ABU08495	ABU08495 Fruit fly
251	7	3.2	561	8	AD066425	AD066425 Aspergillus	324	7	3.2	1074	8	ADJ50367	ADJ50367 Oil-asnoc
252	7	3.2	566	8	AD066691	AD066691 Novel hum	325	7	3.2	1091	6	ABU46052	ABU46052 Protein e
253	7	3.2	568	4	AAW40873	AAW40873 Human pol	326	7	3.2	1091	3	ADK47308	ADK47308 Streptococ
254	7	3.2	594	9	AEA25863	AEA25863 C. elegans	327	7	3.2	1100	3	AAW81511	AAW81511 Streptococ
255	7	3.2	595	4	AAU36514	AAU36514 Pseudomon	328	7	3.2	1100	4	AAW32866	AAW32866 Human pro
256	7	3.2	595	6	ABU38917	ABU38917 Protein e	329	7	3.2	1100	6	ABU01574	ABU01574 S. pneumo
257	7	3.2	597	8	ADN20612	ADN20612 Bacterial	330	7	3.2	1100	6	AAO29563	AAO29563 Human pro
258	7	3.2	602	3	AAW69204	AAW69204 Amino aci	331	7	3.2	1100	9	ADY70491	ADY70491 Human bet
259	7	3.2	603	5	ABO53099	ABO53099 Human put	332	7	3.2	1102	8	ADR96224	ADR96224 Novel S.
260	7	3.2	609	5	ABU65102	ABU65102 Human NOV	333	7	3.2	1102	9	AEA60094	AEA60094 Streptococ
261	7	3.2	609	8	ADN61855	ADN61855 Human nov	334	7	3.2	1106	7	ADW46435	ADW46435 Rat Prote
262	7	3.2	611	4	AAW39087	AAW39087 Human pro	335	7	3.2	1146	7	ADP50150	ADP50150 Murine de
263	7	3.2	621	4	AAW80000	AAW80000 Human pro	336	7	3.2	1147	9	ADW71979	ADW71979 Murine DI
264	7	3.2	628	8	ADN08839	ADN08839 Human DAN	337	7	3.2	1150	5	AAW40294	AAW40294 Human pol
265	7	3.2	639	3	AAW51038	AAW51038 Human pro	338	7	3.2	1150	5	ABG32465	ABG32465 Human pro
266	7	3.2	639	4	AAW78485	AAW78485 Human pro	339	7	3.2	1150	7	ADP50152	ADP50152 Human dea
267	7	3.2	642	3	ABP64400	ABP64400 Human ORF	340	7	3.2	1150	8	ADW88261	ADW88261 Human pro
268	7	3.2	642	5	ABP64400	ABP64400 Human ORF	341	7	3.2	1150	9	ADW71981	ADW71981 Human DIS
269	7	3.2	649	8	ADN24693	ADN24693 Bacterial	342	7	3.2	1159	9	ADY91605	ADY91605 Human pro
270	7	3.2	656	2	AAW04627	AAW04627 Mouse rec	343	7	3.2	1159	9	ADW71952	ADW71952 Murine DI
271	7	3.2	656	2	AAW80994	AAW80994 Human rec	344	7	3.2	1161	7	ADW83142	ADW83142 Murine in
272	7	3.2	656	8	ADJ71949	ADJ71949 Human PMW	345	7	3.2	1162	6	ADW55137	ADW55137 Human pro
273	7	3.2	657	7	ADW48721	ADW48721 Rat Prote	346	7	3.2	1166	3	AAW18268	AAW18268 Plasmodiu
274	7	3.2	661	6	ABP98848	ABP98848 Human str	347	7	3.2	1172	9	ADW71954	ADW71954 Human DIS
275	7	3.2	669	6	ABU38898	ABU38898 Protein e	348	7	3.2	1202	7	ADJ69191	ADJ69191 Human hea
276	7	3.2	669	6	ABU39755	ABU39755 Protein e	349	7	3.2	1227	7	ADW71295	ADW71295 Novel hum
277	7	3.2	673	7	ABW64271	ABW64271 Human pro	350	7	3.2	1233	6	ABP98821	ABP98821 Human str
278	7	3.2	673	7	ABO75176	ABO75176 Pseudomon	351	7	3.2	1258	4	AAW78455	AAW78455 Human pro
279	7	3.2	685	4	AAW79469	AAW79469 Human pro	352	7	3.2	1265	5	ABW57282	ABW57282 Mouse isc
280	7	3.2	685	6	ABR41545	ABR41545 Human DIT	353	7	3.2	1265	8	ADW00989	ADW00989 Mouse hom
281	7	3.2	686	8	ABW84061	ABW84061 Human dia	354	7	3.2	1282	8	ADW21864	ADW21864 Human ves
282	7	3.2	698	3	AAW91524	AAW91524 Human sec	355	7	3.2	1285	6	ABR47540	ABR47540 Human ves
283	7	3.2	698	4	AAW93431	AAW93431 Human pol	356	7	3.2	1285	7	ADW75449	ADW75449 Breast ca
284	7	3.2	698	8	ADL71596	ADL71596 Novel hum	357	7	3.2	1285	8	ADL12778	ADL12778 Human ste
285	7	3.2	698	8	ADL10329	ADL10329 Human pro	358	7	3.2	1285	8	ADW00988	ADW00988 Human hom
286	7	3.2	698	8	ADL10372	ADL10372 Human pro	359	7	3.2	1285	9	ADW71980	ADW71980 Human DIS
287	7	3.2	698	8	ABW80566	ABW80566 Tumour-as	360	7	3.2	1289	9	ADW71978	ADW71978 Murine DI
288	7	3.2	699	6	ABW85513	ABW85513 Human tra	361	7	3.2	1299	3	AAW58633	AAW58633 Protein r
289	7	3.2	701	6	ABU38904	ABU38904 Protein e	362	7	3.2	1299	4	AAW78454	AAW78454 Human pro
290	7	3.2	701	6	ABU41400	ABU41400 Protein e	363	7	3.2	1299	4	ADW09816	ADW09816 Human bre
291	7	3.2	702	8	ADN21371	ADN21371 Protein e	364	7	3.2	1307	4	AAW79438	AAW79438 Human pro
292	7	3.2	702	8	ADN21371	ADN21371 Bacterial	365	7	3.2	1307	4	AAW79439	AAW79439 Human pro
293	7	3.2	703	7	ABO75159	ABO75159 Pseudomon	366	7	3.2	1311	6	ADW10025	ADW10025 Allioiococ
294	7	3.2	705	8	ADJ50107	ADJ50107 Oil-asnoc	367	7	3.2	1315	9	ADW06797	ADW06797 CycLin-de
295	7	3.2	717	8	ADN21501	ADN21501 Bacterial	368	7	3.2	1319	9	ADW07105	ADW07105 CycLin-de
296	7	3.2	718	4	AAW92973	AAW92973 Human pro	369	7	3.2	1336	6	ABW52816	ABW52816 Disease t
297	7	3.2	718	8	ADW88419	ADW88419 Human pro	370	7	3.2	1336	6	ADW62224	ADW62224 Disease t
298	7	3.2	721	5	ADZ26477	ADZ26477 Human ery	371	7	3.2	1336	8	ADW19070	ADW19070 Bacterial
299	7	3.2	730	5	ABW98317	ABW98317 Human leu	372	7	3.2	1383	6	ADW10023	ADW10023 Allioiococ
300	7	3.2	731	4	ABG17452	ABG17452 Novel hum	373	7	3.2	1443	6	ABG74688	ABG74688 Human GGD
301	7	3.2	731	7	ADM03766	ADM03766 Human pro	374	7	3.2	1584	8	ABW60133	ABW60133 Tumour-as
302	7	3.2	731	8	ADW34515	ADW34515 POSH prot	375	7	3.2	1644	7	ADW69113	ADW69113 Human hea
303	7	3.2	743	5	ADQ67542	ADQ67542 Novel hum	376	7	3.2	1898	2	AAW30795	AAW30795 A human c
304	7	3.2	764	5	ADL17007	ADL17007 Bovine NO	377	7	3.2	1898	7	ADW48869	ADW48869 Human pro
305	7	3.2	798	4	AAW93238	AAW93238 Human pro	378	7	3.2	1902	6	ABJ18695	ABJ18695 Human nuc
306	7	3.2	805	7	ABO70119	ABO70119 Pseudomon	379	7	3.2	2000	8	ADN03898	ADN03898 Antipapori
307	7	3.2	857	8	ADW76173	ADW76173 Plant ful	380	7	3.2	2000	8	ADP25346	ADP25346 PRO polyp
308	7	3.2	861	4	ABW63028	ABW63028 Drosophila	381	7	3.2	2285	9	ABW02024	ABW02024 Human MPT
309	7	3.2	871	4	ABW58282	ABW58282 Drosophila	382	7	3.2	2285	9	ABW94379	ABW94379 Human SWI
310	7	3.2	871	5	AAW48928	AAW48928 Thermophi	383	7	3.2	2441	2	AAW79054	AAW79054 CREB bind
311	7	3.2	871	5	AAW48927	AAW48927 Thermophi	384	7	3.2	2441	3	AAW40058	AAW40058 Cellular
312	7	3.2	876	9	AEA25861	AEA25861 C. elegans	385	7	3.2	2441	3	AAW94252	AAW94252 Mouse nuc
313	7	3.2	876	9	AEA25890	AEA25890 C. elegans	386	7	3.2	2441	4	ABW44555	ABW44555 Mouse exp
314	7	3.2	878	7	ADW74139	ADW74139 Human nov	387	7	3.2	2441	6	ABU03977	ABU03977 Human CRE
315	7	3.2	900	7	ADJ69696	ADJ69696 Human hea	388	7	3.2	2441	8	ADP90352	ADP90352 Mouse CRE
316	7	3.2	916	8	ADR89554	ADR89554 Oxyterol	389	7	3.2	2441	8	ADR87292	ADR87292 Mouse CRE

536	6	2.7	56	4	ABB17456	Abb17456 Human ner	609	6	2.7	75	4	AAAM34055	Aam34055 Peptide #
537	6	2.7	56	5	ABG36709	Abg36709 Human pep	610	6	2.7	75	4	ABBB24739	Abb24739 Protein #
538	6	2.7	57	3	AAAG59934	Aag59934 Arabidops	611	6	2.7	75	4	AAAM73873	Aam73873 Human bon
539	6	2.7	58	4	AAU60632	Aau60632 Proportion	612	6	2.7	75	4	AAAM61159	Aam61159 Human bra
540	6	2.7	58	4	AAU63534	Aau63534 Proportion	613	6	2.7	75	4	ABGS5624	Abg55624 Human liv
541	6	2.7	58	6	ABM60053	Abm60053 Proportion	614	6	2.7	75	5	ABG43761	Abg43761 Human pep
542	6	2.7	58	6	ABM65854	Abm65854 Proportion	615	6	2.7	76	3	AAAG55742	Aag555742 Arabidops
543	6	2.7	58	6	ABM57151	Abm57151 Proportion	616	6	2.7	76	3	AAAG60173	Aag60173 Arabidops
544	6	2.7	59	4	AAAG2798	Aag2798 C glutami	617	6	2.7	76	3	AAAG58776	Aag58776 Arabidops
545	6	2.7	59	4	ABG03925	Abg03925 Novel hum	618	6	2.7	76	3	AAAG55353	Aag55353 Arabidops
546	6	2.7	60	3	AAAG15603	Aag15603 Arabidops	619	6	2.7	77	4	AAAM91396	Aam91396 Human imm
547	6	2.7	60	4	AAU50900	Aau50900 Proportion	620	6	2.7	77	2	ABBO80837	Abb080837 pseudomon
548	6	2.7	60	4	AAU46529	Aau46529 Proportion	621	6	2.7	78	2	AAAR15744	Aar15744 R18 antiig
549	6	2.7	60	4	ABG20913	Abg20913 Novel hum	622	6	2.7	78	4	AAAM99828	Aam99828 HIV HBX2
550	6	2.7	60	6	ABM43048	Abm43048 Proportion	623	6	2.7	78	4	AAAB85995	Aab85995 Amino aci
551	6	2.7	60	6	ABM47419	Abm47419 Proportion	624	6	2.7	78	6	AAO30521	Aao30521 HIV vpr m
552	6	2.7	60	7	ABR42471	Abt42471 HIV viral	625	6	2.7	78	6	AAO30515	Aao30515 HIV vpr n
553	6	2.7	61	6	ABM38447	Abm38447 Proportion	626	6	2.7	78	6	AAO30516	Aao30516 HIV vpr m
554	6	2.7	61	6	ABM38447	Abm38447 Proportion	627	6	2.7	78	6	AAO30519	Aao30519 HIV vpr m
555	6	2.7	62	6	ADP35689	Adp35689 Human hep	628	6	2.7	78	6	AAO30517	Aao30517 HIV vpr m
556	6	2.7	62	6	ADP36270	Adp36270 Actinotoba	629	6	2.7	78	6	AAO30518	Aao30518 HIV vpr m
557	6	2.7	62	8	ADP07828	Adp07828 Human sec	630	6	2.7	78	6	AAO30520	Aao30520 HIV vpr m
558	6	2.7	63	4	AAU56121	Aau56121 Proportion	631	6	2.7	78	8	ADN36417	Adn36417 HIV prote
559	6	2.7	63	4	AAU49893	Aau49893 Proportion	632	6	2.7	78	8	ADX96214	Adx96214 Plant ful
560	6	2.7	63	6	ABM46412	Abm46412 Proportion	633	6	2.7	79	2	AAAR48961	Aar48961 NL4-3 VPR
561	6	2.7	63	6	ABM52640	Abm52640 Proportion	634	6	2.7	80	7	ABO60704	Ab060704 Klebsiell
562	6	2.7	64	4	AAU53994	Aau53994 Proportion	635	6	2.7	80	7	ADL27014	Adl27014 Human 339
563	6	2.7	64	6	ABM50513	Abm50513 Proportion	636	6	2.7	81	4	AAU86504	Aau86504 Novel hum
564	6	2.7	65	4	AAU61220	Aau61220 Proportion	637	6	2.7	81	4	AAU50363	Aau50363 Proportion
565	6	2.7	65	6	ABG06896	Abg06896 Novel hum	638	6	2.7	81	5	ABP31762	Abp31762 Human ORF
566	6	2.7	65	6	ABM57739	Abm57739 Proportion	639	6	2.7	81	6	ABM46882	Abm46882 Proportion
567	6	2.7	66	6	ABM71604	Abm71604 Staphyloc	640	6	2.7	81	7	ADBS9838	Adbs9838 Connectiv
568	6	2.7	66	4	AAU66720	Aau66720 Proportion	641	6	2.7	82	4	AAAM83284	Aam83284 Human imm
569	6	2.7	66	4	AAU52198	Aau52198 Proportion	642	6	2.7	82	4	AAU41336	Aau41336 Proportion
570	6	2.7	66	5	AAU60952	Abu60952 Lung spec	643	6	2.7	82	6	ABM37855	Abm37855 Proportion
571	6	2.7	66	5	ADP94731	Adp94731 Hepatitlis	644	6	2.7	83	4	AAU46839	Aau46839 Proportion
572	6	2.7	66	6	ABM48717	Abm48717 Proportion	645	6	2.7	83	4	AAU66297	Aau66297 Proportion
573	6	2.7	66	6	ABM63239	Abm63239 Proportion	646	6	2.7	83	5	ABP42262	Abp42262 Human ova
574	6	2.7	66	7	ADP07640	Adt07640 Bacterial	647	6	2.7	83	5	ABM43358	Abm43358 Proportion
575	6	2.7	67	4	AAAM13788	Aam13788 Peptide #	648	6	2.7	83	6	ABM62816	Abm62816 Proportion
576	6	2.7	67	4	ABM68657	Abm68657 Drosophil	649	6	2.7	83	9	ADX40859	Adx40859 HIV vpr p
577	6	2.7	67	4	ABM41378	Abb41378 Peptide #	650	6	2.7	84	2	AAAG95979	Aag95979 Nucleic a
578	6	2.7	67	4	ABM32722	Abb32722 Peptide #	651	6	2.7	84	5	AAAM52410	Aam52410 Protease
579	6	2.7	67	4	AAAM26186	Aam26186 Peptide #	652	6	2.7	85	3	AAAY64719	Aay64719 Human 5'
580	6	2.7	67	4	AAAM35168	Aam35168 Peptide #	653	6	2.7	85	3	AAAG00616	Aag00616 Human sec
581	6	2.7	67	4	ABM27563	Abb27563 Human pep	654	6	2.7	85	4	ABG05351	Abg05351 Novel hum
582	6	2.7	67	4	ABM18208	Abb18208 Protein #	655	6	2.7	85	5	ABM10002	Abm10002 Human pro
583	6	2.7	67	4	ABM25308	Abb25308 Protein #	656	6	2.7	85	5	ABM10017	Abb10017 Human pro
584	6	2.7	67	4	AAAM75051	Aam75051 Human bon	657	6	2.7	85	5	ABG64636	Abg64636 Human alb
585	6	2.7	67	4	AAAM53540	Aam53540 Human bra	658	6	2.7	85	5	ABG64638	Abg64638 Human alb
586	6	2.7	67	4	AAAM62247	Aam62247 Human bra	659	6	2.7	85	8	ADL77903	Adl77903 Albumin f
587	6	2.7	67	4	ABG47577	Abg47577 Human liv	660	6	2.7	85	8	ADL77905	Adl77905 Albumin f
588	6	2.7	67	4	ABG56816	Abg56816 Human liv	661	6	2.7	85	8	ADU72283	Adu72283 Signal pe
589	6	2.7	67	4	AAAM01533	Aam01533 Peptide #	662	6	2.7	85	9	ADZ73274	Adz73274 Human inc
590	6	2.7	67	5	ABG44757	Abg44757 Human pep	663	6	2.7	86	6	ABU11767	Abu11767 Human MDD
591	6	2.7	68	5	ABG35557	Abg35557 Human pep	664	6	2.7	87	4	AAAM89216	Aam89216 Human imm
592	6	2.7	68	7	ABO66527	Ab066527 Klebsiell	665	6	2.7	88	2	AAAY35878	Aay35878 Amino aci
593	6	2.7	69	6	ABP75975	Abp75975 Human GEN	666	6	2.7	88	3	AAAG44944	Aag44944 Zea may
594	6	2.7	69	6	ABP76141	Abp76141 Human GEN	667	6	2.7	88	4	AAAM95468	Aam95468 Human rep
595	6	2.7	70	4	AAAG75057	Aag75057 Human col	668	6	2.7	88	4	AAU45755	Aau45755 Proportion
596	6	2.7	70	8	ABO54240	Ab054240 Human gen	669	6	2.7	88	4	ABB96153	Abb96153 Human tes
597	6	2.7	71	4	AAAM87393	Aam87393 Human imm	670	6	2.7	88	5	ABP11139	Abp11139 Human ORF
598	6	2.7	71	4	AAAM90859	Aam90859 Human imm	671	6	2.7	88	6	ABM42274	Abm42274 Proportion
599	6	2.7	72	2	AAAR94547	Aar94547 Fragment	672	6	2.7	89	4	ABBB68584	Abb68584 Drosophil
600	6	2.7	72	3	AAAY53250	Aay53250 SIV Vpr p	673	6	2.7	90	3	AAAG44943	Aag44943 Zea may
601	6	2.7	72	3	AAAG61246	Aag61246 Arabidops	674	6	2.7	90	4	AAU50500	Aau50500 Proportion
602	6	2.7	72	3	AAAG55704	Aag55704 Arabidops	675	6	2.7	90	5	ABP04104	Abp04104 Human ORF
603	6	2.7	72	3	AAAG55704	Aag55704 Arabidops	676	6	2.7	90	5	ABP08575	Abp08575 Human ORF
604	6	2.7	72	6	ABM50964	Abm50964 Proportion	677	6	2.7	90	6	ABM47019	Abm47019 Proportion
605	6	2.7	73	3	AAAG44945	Aag44945 Zea may	678	6	2.7	91	3	AAAY65654	Aay65654 C. elegan
606	6	2.7	74	4	ABBB67692	Abb67692 Drosophil	679	6	2.7	91	4	ABBA1436	Abba1436 Peptide #
607	6	2.7	75	4	AAAM20135	Aam20135 Peptide #	680	6	2.7	91	4	AAAM35228	Aam35228 Peptide #
608	6	2.7	75	4	ABBA40351	Abb40351 Peptide #	681	6	2.7	91	4	ABBB25336	Abb25336 Protein #

682	6	2.7	91	4	AAW75112	Aaw75112	Human	bon	755	6	2.7	96	9	ADX40865	Adx40865	HIV	Vpr	P
683	6	2.7	91	4	AAW62309	Aaw62309	Human	bra	756	6	2.7	96	9	ADX40884	Adx40884	HIV	Vpr	P
684	6	2.7	91	4	ABG56875	Abg56875	Human	liv	757	6	2.7	96	9	ADX40891	Adx40891	HIV	Vpr	P
685	6	2.7	91	5	ABG44796	Abg44796	Human	pep	758	6	2.7	96	9	ADX40853	Adx40853	HIV	Vpr	P
686	6	2.7	91	5	ADX40868	Adx40868	HIV	Vpr	759	6	2.7	96	9	ADX40877	Adx40877	HIV	Vpr	P
687	6	2.7	92	3	AAW69305	Aaw69305	HIV-1	non	760	6	2.7	96	9	ADX40877	Adx40877	HIV	Vpr	P
688	6	2.7	92	4	ABG21713	Abg21713	Novel	hum	761	6	2.7	96	9	ADX40844	Adx40844	HIV	Vpr	P
689	6	2.7	92	4	ABG11873	Abg11873	Novel	hum	762	6	2.7	96	9	ADX40873	Adx40873	HIV	Vpr	P
690	6	2.7	92	5	ABP33004	Abp33004	Human	ORF	763	6	2.7	96	9	ADX40881	Adx40881	HIV	Vpr	P
691	6	2.7	93	3	AAW40589	Aaw40589	Human	ORF	764	6	2.7	96	9	ADX40885	Adx40885	HIV	Vpr	P
692	6	2.7	93	3	ABG02968	Abg02968	Novel	hum	765	6	2.7	96	9	ADX40874	Adx40874	HIV-1	Vpr	P
693	6	2.7	93	4	ABG04388	Abg04388	Novel	hum	766	6	2.7	96	9	ADZ71148	Adz71148	HIV-1	ORF	
694	6	2.7	93	4	ABG26512	Abg26512	Novel	hum	767	6	2.7	96	9	ABE10650	Abel10650	Clade C	V	
695	6	2.7	94	5	ABW48618	Abw48618	Listeria		768	6	2.7	96	9	ABE10600	Abel10600	Clade B	V	
696	6	2.7	94	8	ADL90099	Adl90099	Glycoprote		769	6	2.7	96	9	ABE10648	Abel10648	Clade C	V	
697	6	2.7	95	3	AAW69311	Aaw69311	HIV-1	non	770	6	2.7	96	9	ABE10598	Abel10598	Clade B	V	
698	6	2.7	95	3	AAW69307	Aaw69307	HIV-1	non	771	6	2.7	96	9	ABE10599	Abel10599	Clade B	V	
699	6	2.7	95	3	AAW57319	Aaw57319	Propionib		772	6	2.7	96	9	AAU66144	Aau66144	Propionib		
700	6	2.7	95	6	ABW53838	Abw53838	Propionib		773	6	2.7	97	6	ABM62663	Abm62663	Propionib		
701	6	2.7	95	6	ABE78234	Abel78234	Endometri		774	6	2.7	97	6	ABG11735	Abg11735	Novel	hum	
702	6	2.7	95	7	ADL05114	Adl05114	M. catarr		775	6	2.7	98	4	ABP06495	Abp06495	Human	ORF	
703	6	2.7	95	8	ADX40878	Adx40878	HIV	Vpr	776	6	2.7	98	5	ABP05867	Abp05867	Human	ORF	
704	6	2.7	95	9	ADX40850	Adx40850	HIV	Vpr	777	6	2.7	98	5	ABO69858	Abp069858	Pseudomon		
705	6	2.7	96	1	AAW81863	Aaw81863	Sequence		778	6	2.7	98	7	ADX88145	Adx88145	Plant	ful	
706	6	2.7	96	1	AAW12258	Aaw12258	HIV-1	str	779	6	2.7	98	8	AAW01729	Aaw01729	Human	pol	
707	6	2.7	96	2	AAW48963	Aaw48963	HIV	VPR.	780	6	2.7	99	2	AAW05641	Aaw05641	HIV-1	gro	
708	6	2.7	96	2	AAW48944	Aaw48944	Native	VP	781	6	2.7	100	2	ABG44911	Abg44911	Zea	maye	
709	6	2.7	96	2	AAW53040	Aaw53040	HIV-1	pol	782	6	2.7	100	3	ABG424141	Abg424141	Novel	hum	
710	6	2.7	96	2	AAW72995	Aaw72995	HIV	160la	783	6	2.7	101	4	ABR82764	Abbr82764	Novel	14	
711	6	2.7	96	2	AAW29070	Aaw29070	T. gondii		784	6	2.7	101	6	ABR82764	Abbr82764	Novel	14	
712	6	2.7	96	2	AAW99823	Aaw99823	HIV	L685	785	6	2.7	102	4	AAE01795	Aae01795	Human	gen	
713	6	2.7	96	2	AAW99824	Aaw99824	HIV	H71C	786	6	2.7	102	4	ABG05370	Abg05370	Novel	hum	
714	6	2.7	96	2	AAW99821	Aaw99821	HIV	L64S	787	6	2.7	102	5	ABG64168	Abg64168	Enterococ		
715	6	2.7	96	2	AAW99818	Aaw99818	HIV	A30S	788	6	2.7	102	7	ADH87220	Adh87220	Albumin	f	
716	6	2.7	96	2	AAW99826	Aaw99826	HIV	G75A	789	6	2.7	102	8	ADL77433	Adl77433	Propionib		
717	6	2.7	96	2	AAW99812	Aaw99812	HIV-1	Vpr	790	6	2.7	103	4	AAU64024	Aau64024	Streplococ		
718	6	2.7	96	2	AAW99819	Aaw99819	HIV	A30L	791	6	2.7	103	6	ABM60543	Abm60543	Propionib		
719	6	2.7	96	2	AAW99820	Aaw99820	HIV	A59P	792	6	2.7	103	8	ADV81306	Adv81306	Streplococ		
720	6	2.7	96	2	AAW99825	Aaw99825	HIV	H71Y	793	6	2.7	103	8	ADX87863	Adx87863	Plant	ful	
721	6	2.7	96	2	AAW99827	Aaw99827	HIV	C76S	794	6	2.7	104	4	ABR41818	Abbr41818	Peptide	#	
722	6	2.7	96	2	AAW99815	Aaw99815	HIV	Vpr	795	6	2.7	104	4	AAW35614	Aaw35614	Peptide	#	
723	6	2.7	96	2	AAW99829	Aaw99829	HIV	Vpr	796	6	2.7	104	4	AAW75508	Aaw75508	Human	bon	
724	6	2.7	96	2	AAW99822	Aaw99822	HIV	L67S	797	6	2.7	104	4	AAW62690	Aaw62690	Human	bra	
725	6	2.7	96	2	AAW99831	Aaw99831	HIV	E21.2	798	6	2.7	104	4	ABG57256	Abg57256	Human	liv	
726	6	2.7	96	2	AAW99831	Aaw99831	HIV	E21.2	799	6	2.7	104	4	ABG26914	Abg26914	Streplococ		
727	6	2.7	96	2	AAW27492	Aaw27492	E. coli	b	800	6	2.7	104	5	ABP26914	Abp26914	Streplococ		
728	6	2.7	96	3	AAW10049	Aaw10049	HIV-1	Vpr	801	6	2.7	104	8	ADV87854	Adv87854	Streplococ		
729	6	2.7	96	3	AAW69303	Aaw69303	HIV-1	non	802	6	2.7	105	4	ABR03916	Abbr03916	Human	mus	
730	6	2.7	96	3	AAW69306	Aaw69306	HIV-1	non	803	6	2.7	105	5	ABP26915	Abp26915	Streplococ		
731	6	2.7	96	3	AAW69304	Aaw69304	HIV-1	non	804	6	2.7	105	6	ABU13210	Abu13210	Novel	hum	
732	6	2.7	96	3	AAW69302	Aaw69302	HIV-1	reg	805	6	2.7	105	8	ADJ29236	Adj29236	Human	mus	
733	6	2.7	96	3	AAW10685	Aaw10685	HIV-1	reg	806	6	2.7	106	4	ABG00977	Abg00977	Novel	hum	
734	6	2.7	96	3	AAW53247	Aaw53247	HIV-1	LAI	807	6	2.7	106	7	ABM89992	Abm89992	Rice	abio	
735	6	2.7	96	3	AAW53541	Aaw53541	T. gondii		808	6	2.7	107	9	ADZ22236	Adz22236	Full	leng	
736	6	2.7	96	5	AAW61175	Aaw61175	Human	imm	809	6	2.7	107	9	ADZ22234	Adz22234	Full	leng	
737	6	2.7	96	5	AAW61175	Aaw61175	HIV-1	vir	810	6	2.7	107	9	ABE07672	Abel07672	PDZ	domai	
738	6	2.7	96	5	AAW61175	Aaw61175	HIV-1	vir	811	6	2.7	107	9	ABE07672	Abel07672	PDZ	domai	
739	6	2.7	96	5	AAW61175	Aaw61175	HIV-1	vir	812	6	2.7	107	9	ABE07672	Abel07672	PDZ	domai	
740	6	2.7	96	7	ADG17316	Adg17316	T. gondii		813	6	2.7	108	6	ABM46530	Abm46530	Full	leng	
741	6	2.7	96	8	ADP46812	Adp46812	HIV-1	Vpr	814	6	2.7	109	3	AAW02979	Aaw02979	Human	sec	
742	6	2.7	96	8	ABW79669	Abw79669	HIV-1	Vpr	815	6	2.7	109	3	AAW02979	Aaw02979	Human	sec	
743	6	2.7	96	8	ADP052552	Adp052552	HIV-1	reg	816	6	2.7	109	4	ABG06515	Abg06515	Novel	hum	
744	6	2.7	96	8	ADP052552	Adp052552	HIV-1	reg	817	6	2.7	109	4	ABG06515	Abg06515	Novel	hum	
745	6	2.7	96	8	ADP052552	Adp052552	Human	imm	818	6	2.7	109	5	ABP09241	Abp09241	Human	ORF	
746	6	2.7	96	8	ADP052552	Adp052552	Human	imm	819	6	2.7	109	5	ABP09241	Abp09241	Human	ORF	
747	6	2.7	96	9	ADV23774	Adv23774	HIV-1	ful	820	6	2.7	109	7	ADP31761	Adp31761	Human	nov	
748	6	2.7	96	9	ADX40841	Adx40841	HIV	Vpr	821	6	2.7	109	8	ADP31761	Adp31761	Human	nov	
749	6	2.7	96	9	ADX40869	Adx40869	HIV	Vpr	822	6	2.7	109	8	ADP31761	Adp31761	Human	nov	
750	6	2.7	96	9	ADX40849	Adx40849	HIV	Vpr	823	6	2.7	109	8	ADP31761	Adp31761	Human	nov	
751	6	2.7	96	9	ADX40845	Adx40845	HIV	Vpr	824	6	2.7	109	8	ADP31761	Adp31761	Human	nov	
752	6	2.7	96	9	ADX40847	Adx40847	HIV	Vpr	825	6	2.7	109	8	ADP31761	Adp31761	Human	nov	
753	6	2.7	96	9	ADX40854	Adx40854	HIV	Vpr	826	6	2.7	109	8	ADP31761	Adp31761	Human	nov	
754	6	2.7	96	9	ADX40864	Adx40864	HIV	Vpr	827	6	2.7	109	9	ADP31761	Adp31761	Human	nov	

828	2.7	110	4	AAM90884	Human	Imm	901	6	2.7	125	4	ABB20920	Abb20920	Protein #
829	2.7	110	4	ABE6519	Human	NR	902	6	2.7	125	4	AAM68679	Aam68679	Human bon
830	2.7	110	5	ABP09182	Human	ORF	903	6	2.7	125	4	AAW56301	Aaw56301	Human bra
831	2.7	110	2	ADY37612	Lung	canc	904	6	2.7	125	4	ABG50343	Abg50343	Human liv
832	2.7	111	2	AAM38667	Streptoco		905	6	2.7	125	4	AAM04220	Aam04220	Peptide #
833	2.7	111	4	AAM84540	Human	Imm	906	6	2.7	125	5	ABG38258	Abg38258	Human pep
834	2.7	111	4	AAU48005	Proponib		907	6	2.7	126	3	AAAG4910	Aag4910	Zea maye
835	2.7	111	6	ABM44524	Proponib		908	6	2.7	126	4	ABG79706	Abg79706	Coryneb
836	2.7	112	8	ADN17279	Arabidops		909	6	2.7	126	7	ABG72980	Abg72980	Pseudom
837	2.7	112	8	ADR16842	Arabidops		910	6	2.7	127	7	ABG20770	Abg20770	Novel hum
838	2.7	112	9	ADV66224	Arabidops		911	6	2.7	127	8	ADSI15037	Adsi15037	Pseudom
839	2.7	113	3	AAW25179	Eucalectu		912	6	2.7	127	8	ADW95363	Adw95363	Plant ful
840	2.7	113	4	AAU52411	Proponib		913	6	2.7	128	2	AAM88539	Aam88539	Secreted
841	2.7	113	6	ABM48930	Proponib		914	6	2.7	128	4	ABW50306	Abw50306	Human sec
842	2.7	113	6	ABU43568	Proponib		915	6	2.7	128	5	ADK35127	Adk35127	Novel hum
843	2.7	113	6	ABU42565	Protein e		916	6	2.7	128	6	ABO44563	AbO44563	Novel hum
844	2.7	113	6	ABU16015	Protein e		917	6	2.7	128	7	ABO26043	AbO26043	Human pro
845	2.7	113	7	ADU12429	PDZ	ligan	918	6	2.7	129	7	AAW75924	Aaw75924	Human pol
846	2.7	113	8	AD153451	Human	PDZ	919	6	2.7	129	4	AAO01201	Aao01201	Human pol
847	2.7	113	8	ADM33542	Human	PDZ	920	6	2.7	129	7	ADC37544	Adc37544	Human nuc
848	2.7	113	8	ADO20973	PDZ	domai	921	6	2.7	129	8	ADW96651	Adw96651	Plant ful
849	2.7	113	8	ADR82914	PDZ	domai	922	6	2.7	130	2	AAV42116	Aav42116	Soybean a
850	2.7	113	8	ADU15906	Human	KIA	923	6	2.7	130	4	AAO05077	Aao05077	Human pol
851	2.7	113	8	ADU67331	Human	KIA	924	6	2.7	130	4	AAU30652	Aau30652	Novel hum
852	2.7	113	9	ADW52545	Human	PDZ	925	6	2.7	130	5	ABR40527	AbR40527	Human sec
853	2.7	113	9	AEB07696	PDZ	domai	926	6	2.7	130	5	ABR40448	AbR40448	Human sec
854	2.7	114	9	AEA98710	Human	PDZ	927	6	2.7	131	3	AAAG10065	Aag10065	Arabidops
855	2.7	114	4	ABBI2260	Human	sec	928	6	2.7	131	3	AAAG37424	Aag37424	Arabidops
856	2.7	114	4	AAAM0066	Human	pro	929	6	2.7	131	4	AAU48942	Aau48942	Proponib
857	2.7	115	4	AAU21302	Human	nov	930	6	2.7	131	6	ABM45461	Abm45461	Proponib
858	2.7	116	5	ABG29960	Novel	hum	931	6	2.7	131	8	ADT55960	Adt55960	Plant pol
859	2.7	116	5	AAE25941	Soybean	G	932	6	2.7	132	2	AAW55365	Aaw55365	H. pylori
860	2.7	116	6	ABU67338	Soybean	G	933	6	2.7	132	3	AAAG10051	Aag10051	Human sec
861	2.7	116	8	ADY78257	Plant	ful	934	6	2.7	132	3	AAAG36582	Aag36582	Arabidops
862	2.7	116	8	ABE11240	Soybean	G	935	6	2.7	132	4	ABBI1003	Abbi1003	Human sec
863	2.7	116	9	ABE48361	Soybean	G	936	6	2.7	134	5	ABP40764	AbP40764	Staphyloc
864	2.7	117	2	AAW28002	Amino	acti	937	6	2.7	134	5	ABBS7221	Abbs7221	Mouse
865	2.7	117	4	AAU36840	Staphyloc		938	6	2.7	134	7	ADBE2496	Adbe2496	Human pro
866	2.7	117	4	AAU37274	Staphyloc		939	6	2.7	134	7	ADBE2494	Adbe2494	Rat
867	2.7	117	4	AAU37501	Staphyloc		940	6	2.7	134	7	ABO78213	AbO78213	Pseudom
868	2.7	117	6	ABM72404	Staphyloc		941	6	2.7	134	8	ADSO8081	Adso8081	Staphyloc
869	2.7	117	7	ADD35694	Human	hep	942	6	2.7	134	8	ADX68658	Adx68658	Plant ful
870	2.7	118	4	AAAG74149	Human	col	943	6	2.7	134	9	ABE91378	AbE91378	Microbia
871	2.7	119	3	AAI13260	Human	met	944	6	2.7	135	5	ABE89058	AbE89058	Human pol
872	2.7	119	4	ABBI5306	Human	ner	945	6	2.7	135	5	ABP32977	AbP32977	Human ORF
873	2.7	119	5	ABP35342	Human	ORF	946	6	2.7	135	8	ADY06637	Ady06637	Plant ful
874	2.7	119	8	ADT58405	Plant	pol	947	6	2.7	135	8	ADW95699	Adw95699	Plant ful
875	2.7	120	2	AAW52825	Human	TPE	948	6	2.7	136	4	AAU43396	Aau43396	Proponib
876	2.7	120	4	AAU61763	Proponib		949	6	2.7	136	6	ABM39915	Abm39915	Proponib
877	2.7	120	7	ABM89852	Proponib		950	6	2.7	136	8	ADX94121	Adx94121	Plant ful
878	2.7	120	7	ABM89852	Rice	abio	951	6	2.7	137	4	AAW85004	Aaw85004	Shrimp wh
879	2.7	121	8	ADW86465	Aspergill		952	6	2.7	137	5	ADG79542	Adg79542	Human sec
880	2.7	122	4	AAU40049	Proponib		953	6	2.7	137	8	ADY11590	Ady11590	Plant ful
881	2.7	122	4	AAU64680	Proponib		954	6	2.7	138	6	ABR43148	AbR43148	Ryegrass
882	2.7	122	6	AAU67329	Proponib		955	6	2.7	138	7	ABO70911	AbO70911	Pseudom
883	2.7	122	6	ABM63848	Proponib		956	6	2.7	139	7	ABG09938	AbG09938	Novel hum
884	2.7	122	6	ABM61199	Proponib		957	6	2.7	139	7	ABO73918	AbO73918	Pseudom
885	2.7	122	6	ABM36568	Proponib		958	6	2.7	139	8	ADW72208	Adw72208	Plant ful
886	2.7	122	8	ADJ46391	B. subcili		959	6	2.7	140	3	AAAG4909	Aag4909	Zea maye
887	2.7	122	8	ADQ089359	Mouse	ser	960	6	2.7	140	3	AAW79392	Aaw79392	Human pro
888	2.7	123	3	AAW96514	P. furios		961	6	2.7	140	7	ADP06215	Adp06215	Bacterial
889	2.7	123	4	ABH11154	Human	sec	962	6	2.7	140	8	ADW71124	Adw71124	Plant ful
890	2.7	124	3	AAW25953	Zea	maye	963	6	2.7	141	3	AAAG10064	Aag10064	Arabidops
891	2.7	124	4	AAW88443	Human	Imm	964	6	2.7	141	3	AAAG37423	Aag37423	Arabidops
892	2.7	124	5	AAO03526	Staphyloc		965	6	2.7	141	6	ADAS5626	Ada55626	Human sig
893	2.7	124	5	ABP40439	Staphyloc		966	6	2.7	141	6	ABU61051	AbU61051	Human ova
894	2.7	124	8	ADSO6053	Staphyloc		967	6	2.7	141	6	ABU61051	AbU61051	Human ova
895	2.7	125	3	AAW89798	Plant	ful	968	6	2.7	142	4	ABBB94189	AbB94189	Human nov
896	2.7	125	3	AAW40767	Human	ORF	969	6	2.7	142	4	ABBB6675	AbB6675	Drosophi
897	2.7	125	4	AAW16497	Peptide	#	970	6	2.7	142	4	AAW062638	Aaw062638	Novel hum
898	2.7	125	4	ABW35483	Peptide	#	971	6	2.7	142	4	ABG05888	AbG05888	Novel hum
899	2.7	125	4	AAW28985	Peptide	#	972	6	2.7	142	5	ABP35032	AbP35032	Human ORF
900	2.7	125	4	ABW30311	Peptide	#	973	6	2.7	142	6	ABW59157	AbW59157	Proponib

974	6	2.7	142	8	AD142470	Ad142470 Plant tra
975	6	2.7	143	3	AA93242	AA93242 An Escher
976	6	2.7	143	7	ADH87764	ADH87764 Enterococ
977	6	2.7	143	8	ADO22037	ADO22037 Silicibac
978	6	2.7	143	8	ADY05647	ADY05647 Plant ful
979	6	2.7	143	9	ABE29069	ABE29069 Human OVO
980	6	2.7	143	9	ABE29067	ABE29067 Human OVO
981	6	2.7	143	9	ABE25420	ABE25420 Pinus rad
982	6	2.7	145	3	AB43074	AB43074 Human ORF
983	6	2.7	146	4	ABG20929	ABG20929 Novel hum
984	6	2.7	146	5	ADK35186	ADK35186 Novel hum
985	6	2.7	146	6	ABU33774	ABU33774 Protein e
986	6	2.7	147	6	ABP60617	ABP60617 Human hum
987	6	2.7	148	3	AA945233	AA945233 Arabidops
988	6	2.7	148	7	AB077145	AB077145 Pseudomon
989	6	2.7	149	3	AA917619	AA917619 Arabidops
990	6	2.7	149	4	ABG07326	ABG07326 Novel hum
991	6	2.7	149	4	ABG29930	ABG29930 Novel hum
992	6	2.7	149	7	ADH84387	ADH84387 MSRV-1 as
993	6	2.7	149	7	ADM26145	ADM26145 Hyperther
994	6	2.7	149	8	ADG14833	ADG14833 MSRV asso
995	6	2.7	151	2	AA959910	AA959910 Opacity a
996	6	2.7	151	3	AA948037	AA948037 Arabidops
997	6	2.7	151	3	AA918704	AA918704 Arabidops
998	6	2.7	151	3	AA932302	AA932302 Arabidops
999	6	2.7	151	8	ADO62119	ADO62119 Transcrip
1000	6	2.7	152	3	AA956429	AA956429 Human pro

ALIGNMENTS

RESULT 1
ADPF74494
ID ADPF74494 standard; protein; 220 AA.

AC ADPF74494;

DT 26-FEB-2004 (first entry)

DE Murine elf-3 protein used to treat end stage liver disease.

XX murine; mouse; early liver development; end stage liver disease; elf 1-3;
XX mylor-1 (145); pk; protein 106; praja-1; hepatocyte lineage; cholestasis;
XX biliary stone; hepatic obstruction; stricture; primary biliary cirrhosis;
XX primary sclerosing cholangitis; gene therapy;
XX anhidrotic ectoderm dysplasia; hepatocellular carcinoma; anaemia; ataxia;
XX neurological disorder; haemochromatosis; hepatotropic; cytosstatic;
XX neuroprotective; anti-naemic; cardiant.

OS Mus musculus.

XX US6642362-B1.

PN 04-NOV-2003.

XX 01-NOV-1999; 99US-00431184.

PF 30-APR-1997; 97US-00841349.

PR 30-APR-1998; 98WO-US008656.

XX (MISH/) MISHRA L.

XX Mishra L;

XX WPI, 2003-851362/79.

XX N-PSDB; ADPF74493.

XX New antibodies recognizing early liver development proteins, useful as
PT markers, in identifying peptides and proteins having early liver
PT development characteristics, tracing hepatocyte lineage or treating liver
PT disease.

PS Disclosure; SEQ ID NO 7; 82pp; English.
XX This invention relates to novel genes and encoded proteins thereof.
XX isolated during early liver development that are useful in the diagnosis
XX and treatment of end stage liver disease and other disorders.
XX Specifically, it refers to genes that encode proteins such as elf 1-3,
XX mylor-1 (145), pk, protein 106 and praja-1. The present invention
XX describes the characterisation of these early liver development proteins,
XX and also methods to raise peptide specific antibodies that are useful as
XX markers, as well as for tracing hepatocyte lineage. Furthermore, elf
XX proteins 1-3 are useful in treating disorders including cholestasis,
XX biliary stones, hepatic obstruction, stricture, primary biliary cirrhosis
XX or primary sclerosing cholangitis. In addition, through using gene
XX therapy, praja-1, mylor-1 (145) and pk can be used to treat anhidrotic
XX ectoderm dysplasia, hepatocellular carcinoma and other diseases including
XX anaemia, ataxia, degenerative neurological disorders and
XX haemochromatosis. Accordingly, these proteins can be described as
XX hepatotropic, cytosstatic, neuroprotective, anti-naemic and cardiant. This
XX polypeptide sequence is the murine elf-3 protein of the invention.

SQ Sequence 220 AA;

Query Match 100.0%; Score 220; DB 7; Length 220;
Beet Local Similarity 100.0%; Pred. No. 4,7e-208;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELQRTSSISGPLSPAYTGVPPYNNQLEGRFKQLQDEREAVQKTFKMNSHLARVSC 60
DB 1 MELQRTSSISGPLSPAYTGVPPYNNQLEGRFKQLQDEREAVQKTFKMNSHLARVSC 60
QY 61 RTTDLTYLDRGRLIKLEVLGSRRLPKTKGRIRICLENVDKALFLKEORHLNNM 120
DB 61 RTTDLTYLDRGRLIKLEVLGSRRLPKTKGRIRICLENVDKALFLKEORHLNNM 120
QY 121 GSHDIVDGNHRLTLELEVRROQEEERKCRPPSPDPNTKYSEAEQQMDTSKQDQVS 180
DB 121 GSHDIVDGNHRLTLELEVRROQEEERKCRPPSPDPNTKYSEAEQQMDTSKQDQVS 180
QY 181 QNGLPABQSPRVSYRSQTYQNYKNFNSRRRTASDHSWGM 220
DB 181 QNGLPABQSPRVSYRSQTYQNYKNFNSRRRTASDHSWGM 220

RESULT 2
ADQ88366
ID ADQ88366 standard; protein; 220 AA.

AC ADQ88366;

DT 07-OCT-2004 (first entry)

DE Mouse elf-3 protein.

XX Early liver developmental protein; liver disease;
XX hepatocellular carcinoma; anhidrotic ectoderm dysplasia;
XX degenerative neurological disorder; anaemia; ataxia; haemochromatosis;
XX sideroblastic anaemia; spinocerebellar ataxia; cholestasis;
XX biliary stone; liver obstruction; stricture; primary biliary cirrhosis;
XX primary sclerosing cholangitis; colon cancer; therapy; mouse; elf-1;
XX embryonic liver; fodrin; beta spectrin.

OS Mus musculus.

XX US2004142354-A1.

XX 22-JUL-2004.

XX 30-OCT-2003; 2003US-00695994.

XX 30-APR-1997; 97US-00841349.

XX 30-APR-1998; 98WO-US008656.

XX 01-NOV-1999; 99US-00431184.

PA (MISH/) MISHRA L.
 XX
 XX Mishra L;
 XX
 DR WPI, 2004-552558/53.
 DR N-PSDB; ADO88365.
 XX
 PT Novel isolated early liver developmental proteins such as elf, liver-1
 PT (145), pk, protein 106 and praja-1, useful for treating end stage liver
 PT disease, hepatocellular carcinoma, anemia and ataxia.
 XX
 PS Claim 2; SEQ ID NO 7; 85bp; English.
 XX
 CC The invention provides novel early liver developmental proteins such as
 CC elf (embryonic liver fodrin or beta spectrin) 1-3, liver-1 (145), pk,
 CC protein 106 and praja-1 and nucleic acid molecules encoding such
 CC proteins. Other genes coding for early liver developmental proteins
 CC include genes 20, 36, 41, 112, 114, 118 and 129. Polynucleotides of the
 CC invention are useful for treating diseases chosen from end stage liver
 CC disease, hepatocellular carcinoma, anhidrotic ectoderm dysplasia,
 CC degenerative neurological disorders, anaemia, ataxia, haemochromatosis,
 CC sideroblastic anemia and spinocerebellar ataxia. elf sequences are
 CC useful for treating liver disorders chosen from cholestasis, biliary
 CC stones, liver obstruction, stricture, primary biliary cirrhosis and
 CC primary sclerosing cholangitis. praja-1 sequences are useful for
 CC detecting colon cancer. The present sequence is mouse early liver
 CC developmental protein.
 CC
 XX
 SQ Sequence 220 AA;

Query Match 100.0%; Score 220; DB 8; Length 220;
 Best Local Similarity 100.0%; Pred. No. 4.7e-208;
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELQRTSSISGPLSPAYTGQVPYNNQLEGRFKQLODEREAVOKKFTKWNSHLARVSC 60
 DB 1 MELQRTSSISGPLSPAYTGQVPYNNQLEGRFKQLODEREAVOKKFTKWNSHLARVSC 60
 QY 61 RTDLYTLRDRGMILKLEVLSEGRLPKPTKGRMRHCLENVDKALQFLKEQRYHLENM 120
 DB 61 RTDLYTLRDRGMILKLEVLSEGRLPKPTKGRMRHCLENVDKALQFLKEQRYHLENM 120
 QY 121 GSHDIVDGNHRLTTELEVRQOEERKRKRPSPDPNTKVSSEAESQOMDTSKGDVYS 180
 DB 121 GSHDIVDGNHRLTTELEVRQOEERKRKRPSPDPNTKVSSEAESQOMDTSKGDVYS 180
 QY 181 QNGLPABEQSPRVSYRSQTYQNYKNFNSRRTASDHSWSGM 220
 DB 181 QNGLPABEQSPRVSYRSQTYQNYKNFNSRRTASDHSWSGM 220

RESULT 3
 AAW81640
 ID AAW81640 standard; protein; 220 AA.
 XX
 AC AAW81640;
 XX
 DT 01-MAR-1999 (first entry)
 XX
 DE Mouse elf-3 protein.
 XX
 KW Elf-3; mouse; liver development; signal transduction; biliary tree;
 KW cholestasis; biliary stone; hepatic obstruction; stricture;
 KW primary biliary cirrhosis; primary sclerosing cholangitis; therapy.
 XX
 OS Mus sp.
 XX
 PN WO9848827-A1.
 XX
 PD 05-NOV-1998.
 XX
 PF 30-APR-1998; 98WO-US008656.
 XX

PR 30-APR-1997; 97US-00841349.
 XX
 XX (MISH/) MISHRA L.
 XX
 XX Mishra L;
 XX
 DR WPI, 1999-009382/01.
 DR N-PSDB; AAV64413.
 XX
 PT New isolated early liver development genes - used to develop products for
 PT treating, e.g. liver disease, hepatocellular carcinoma, degenerative
 PT neurological disorders, anaemia, ataxia or haemochromatosis.
 XX
 PS Claim 9; Fig 2c; 92bp; English.
 XX

CC This is the amino acid sequence of the elf-3 protein of the foetal murine
 CC liver. The invention provides early developing liver proteins (see
 CC AAW81638-42) and the genes coding for them (see AAV64410-24). The genes
 CC were isolated from embryonic cDNA libraries using subtractive
 CC hybridisation. 3 Elf genes (see AAV64411-13) were identified. The elf
 CC proteins (see also AAW81639) are probably important for the formation of
 CC the biliary tree during early liver development. They are used in a
 CC claimed method to treat cholestasis, biliary stones, liver obstruction,
 CC stricture, primary biliary cirrhosis and primary sclerosing cholangitis.
 CC Early developing liver proteins and nucleic acids can also be used in the
 CC diagnosis of liver diseases and other disorders, including those relating
 CC to oncogenesis and tissue repair
 CC
 XX
 SQ Sequence 220 AA;

Query Match 58.6%; Score 129; DB 2; Length 220;
 Best Local Similarity 100.0%; Pred. No. 2.2e-118;
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELQRTSSISGPLSPAYTGQVPYNNQLEGRFKQLODEREAVOKKFTKWNSHLARVSC 60
 DB 1 MELQRTSSISGPLSPAYTGQVPYNNQLEGRFKQLODEREAVOKKFTKWNSHLARVSC 60
 QY 61 RTDLYTLRDRGMILKLEVLSEGRLPKPTKGRMRHCLENVDKALQFLKEQRYHLENM 120
 DB 61 RTDLYTLRDRGMILKLEVLSEGRLPKPTKGRMRHCLENVDKALQFLKEQRYHLENM 120
 QY 121 GSHDIVDGN 129
 DB 121 GSHDIVDGN 129

RESULT 4
 ADF13682
 ID ADF13682 standard; protein; 1022 AA.
 XX
 AC ADF13682;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Mouse beta II spectrin amino acid sequence #SEQ ID 2.
 XX
 KW Gene therapy; Sjogren's Syndrome; beta II spectrin.
 XX
 OS Mus sp.
 XX
 PN WO2003100426-A1.
 XX
 PD 04-DEC-2003.
 XX
 PF 20-MAY-2003; 2003WO-US016103.
 XX
 PR 20-MAY-2002; 2002US-0381899P.
 XX
 PA (UYWR-) UNIV WRIGHT STATE.
 PA (UYVA) UNIV YALE.
 XX
 PI Brown TL, Morrow JS;

XX WPI, 2004-062080/06.
 DR N-PEDB; ADF13681.
 XX
 PT Identifying primary and secondary Sjogren's Syndrome or susceptibility to
 PT developing the disease by contacting the sample with a target antigen and
 PT detecting the presence of autoantibodies in the sample that bind to the
 PT target antigen.
 PS Disclosure; SEQ ID NO 2; 40bp; English.
 XX
 CC The invention relates to a method for identifying primary and secondary
 CC Sjogren's Syndrome, or a susceptibility to developing primary and
 CC secondary Sjogren's Syndrome. The method comprises obtaining a biological
 CC sample from the subject, contacting the sample with a purified target
 CC antigen, where the target antigen is an epitope of beta II spectrin and
 CC detecting the presence of autoantibodies in the sample which bind to the
 CC target antigen. The method is useful for identifying primary and
 CC secondary Sjogren's Syndrome or susceptibility to developing primary
 CC pathological effects of primary and secondary Sjogren's Syndrome. The
 CC current sequence represents the mouse beta II spectrin amino acid
 CC sequence.
 XX
 SQ Sequence 1022 AA;

Query Match 56.4%; Score 124; DB 8; Length 1022;
 Best Local Similarity 100.0%; Pred. No. 7e-113;
 Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SGPSPATYGVPPYNNQLEGRFKQLODEREAVOKKFTKWNVSHLAVSCRTIDLYTDL 69
 DB 10 SGPSPATYGVPPYNNQLEGRFKQLODEREAVOKKFTKWNVSHLAVSCRTIDLYTDL 69
 QY 70 RGGRLIKLLEVLGGERLPKPTKGRMRHICLENVDKALQFLKEQRYHLENNGSHDIVDGN 129
 DB 70 RGGRLIKLLEVLGGERLPKPTKGRMRHICLENVDKALQFLKEQRYHLENNGSHDIVDGN 129
 QY 130 HRLT 133
 DB 130 HRLT 133

RESULT 5
 AAW81639
 ID AAW81639 standard; protein; 2154 AA.

XX AC AAW81639;
 XX
 DT 01-MAR-1999 (first entry)
 XX
 DE Mouse elf-1 protein.
 XX
 KW Elf-1; mouse; liver development; signal transduction; biliary tree;
 KW cholestasis; biliary stone; hepatic obstruction; stricture;
 KW primary biliary cirrhosis; primary sclerosing cholangitis; therapy.
 XX
 OS Mus sp.
 XX
 PN WO9848827-A1.
 XX
 PD 05-NOV-1998.
 XX
 PF 30-APR-1998; 98WO-US008656.
 XX
 PR 30-APR-1997; 97US-00841349.
 XX
 PA (MISH/) MISHRA L.
 XX
 PI Mishra L;
 XX
 DR WPI, 1999-009382/01.
 DR N-PEDB; AAV64411.

XX New isolated early liver development genes - used to develop products for
 PT treating, e.g. liver disease, hepatocellular carcinoma, degenerative
 PT neurological disorders, anaemia, ataxia or haemochromatosis.
 PS Claim 9; Fig 2a; 92pp; English.

XX This is the amino acid sequence of the elf-1 protein of the developing
 CC murine foetal liver. The invention provides early developing liver
 CC proteins (see AAW81638-42) and the genes coding for them (see AAV64410-
 CC 24). The genes were isolated from embryonic cDNA libraries using
 CC subtractive hybridisation. 3 Elf genes (see AAV64411-13) were identified.
 CC The elf proteins (see also AAW81640) are probably important for the
 CC formation of the biliary tree during early liver development. They are
 CC used in a claimed method to treat cholestasis, biliary stones, liver
 CC obstruction, stricture, primary biliary cirrhosis and primary sclerosing
 CC cholangitis. Early developing liver proteins and nucleic acids can also
 CC be used in the diagnosis of liver diseases and other disorders, including
 CC those relating to oncogenesis and tissue repair

SQ Sequence 2154 AA;

Query Match 56.4%; Score 124; DB 2; Length 2154;
 Best Local Similarity 100.0%; Pred. No. 1.3e-112;
 Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SGPSPATYGVPPYNNQLEGRFKQLODEREAVOKKFTKWNVSHLAVSCRTIDLYTDL 69
 DB 10 SGPSPATYGVPPYNNQLEGRFKQLODEREAVOKKFTKWNVSHLAVSCRTIDLYTDL 69
 QY 70 RGGRLIKLLEVLGGERLPKPTKGRMRHICLENVDKALQFLKEQRYHLENNGSHDIVDGN 129
 DB 70 RGGRLIKLLEVLGGERLPKPTKGRMRHICLENVDKALQFLKEQRYHLENNGSHDIVDGN 129
 QY 130 HRLT 133
 DB 130 HRLT 133

RESULT 6
 ADF74491
 ID ADF74491 standard; protein; 2154 AA.

XX ADF74491;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Murine elf-1 protein used to treat end stage liver disease.
 XX
 KW murine; mouse; early liver development; end stage liver disease; elf 1-3;
 KW liver-1 (145); pk; protein 106; praja-1; hepatocyte lineage; cholestasis;
 KW biliary stone; hepatic obstruction; stricture; primary biliary cirrhosis;
 KW primary sclerosing cholangitis; gene therapy;
 KW anidrotic ectoderm dysplasia; hepatocellular carcinoma; anaemia; ataxia;
 KW neurological disorder; haemochromatosis; hepatotropic; cytostatic;
 KW neuroprotective; anti-naemic; cardiant.
 XX
 OS Mus musculus.
 XX
 PN US6642362-B1.
 XX
 PD 04-NOV-2003.
 XX
 PF 01-NOV-1999; 99US-00431184.
 XX
 PR 30-APR-1997; 97US-00841349.
 XX
 PR 30-APR-1998; 98WO-US008656.
 XX
 PA (MISH/) MISHRA L.
 XX
 PI Mishra L;
 XX
 DR WPI, 2003-851362/79.

DR N-PSDB; ADF74490.
XX New antibodies recognizing early liver development proteins, useful as
PT markers, in identifying peptides and proteins having early liver
PT development characteristics, tracing hepatocyte lineage or treating liver
PT disease.
XX
XX
PS Disclosure; SEQ ID NO 4; 82pp; English.
XX
CC This invention relates to novel genes and encoded proteins thereof,
CC isolated during early liver development that are useful in the diagnosis
CC and treatment of end stage liver disease and other disorders.
CC Specifically, it refers to genes that encode proteins such as elf 1-3,
CC 11yor-1 (145), pk, protein 106 and praja-1. The present invention
CC describes the characterisation of these early liver development proteins,
CC and also methods to raise peptide specific antibodies that are useful as
CC markers, as well as for tracing hepatocyte lineage. Furthermore, elf
CC proteins 1-3 are useful in treating disorders including cholestasis,
CC biliary stones, hepatic obstruction, stricture, primary biliary cirrhosis
CC or primary sclerosing cholangitis. In addition, through using gene
CC therapy, praja-1, 11yor-1 (145) and pk can be used to treat anhidrotic
CC ectoderm dysplasia, hepatocellular carcinoma and other diseases including
CC anaemia, ataxia, degenerative neurological disorders and
CC haemochromatosis. Accordingly, these proteins can be described as
CC hepatotropic, cytosolic, neuroprotective, antianemic and cardiant. This
CC polypeptide sequence is the murine elf-1 protein of the invention.
XX
SQ Sequence 2154 AA;

Query Match 56.4%; Score 124; DB 7; Length 2154;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SGSPSPATGVPPVNNOLBGRFLODERAVQKTTTKVNSHLAVSCRITDLYTDL 69
Db 10 SGSPSPATGVPPVNNOLBGRFLODERAVQKTTTKVNSHLAVSCRITDLYTDL 69
Db 70 RDGRMLIKLLEVLSEGERLPKPTKGRMRHCLLENVDKALQFLKEQVHLENNGSHDIVDGN 129
QY 130 HRLT 133
Db 70 RDGRMLIKLLEVLSEGERLPKPTKGRMRHCLLENVDKALQFLKEQVHLENNGSHDIVDGN 129
QY 130 HRLT 133
Db 130 HRLT 133

RESULT 7
ADQ88363
ID ADQ88363 standard; protein; 2154 AA.
XX
AC ADQ88363;
XX
DT 07-OCT-2004 (first entry)
XX
DE Mouse elf-1 protein.
XX
XX Early liver developmental protein; liver disease;
KM hepatocellular carcinoma; anhidrotic ectoderm dysplasia;
KM degenerative neurological disorder; anaemia; ataxia; haemochromatosis;
KM sideroblastic anaemia; spinocerebellar ataxia; cholestasis;
KM biliary stone; liver obstruction; stricture; primary biliary cirrhosis;
KM primary sclerosing cholangitis; colon cancer; therapy; mouse; elf-1;
KM embryonic liver fodrin; beta spectrin.
XX
XX Mus musculus.
OS
XX US2004142354-A1.
XX
XX 22-JUL-2004.
XX
XX 30-OCT-2003; 2003US-00695994.
XX
XX 30-APR-1997; 97US-00841349.
PR

PR 30-APR-1998; 98WO-US008656.
PR 01-NOV-1999; 99US-00431184.
XX
XX (MISH/) MISHRA L.
PA
XX
XX Misha L;
PI
XX
XX WPI; 2004-552558/53.
DR
DR N-PSDB; ADQ88362.
XX
XX
XX Novel isolated early liver developmental proteins such as elf, 11yor-1
PT (145), pk, protein 106 and praja-1, useful for treating end stage liver
PT disease, hepatocellular carcinoma, anemia and ataxia.
XX
XX
PS Disclosure; SEQ ID NO 4; 85pp; English.
XX
XX The invention provides novel early liver developmental proteins such as
CC elf (embryonic liver fodrin or beta spectrin) 1-3, 11yor-1 (145), pk,
CC protein 106 and praja-1 and nucleic acid molecules encoding such
CC proteins. Other genes coding for early liver developmental proteins
CC include genes 20, 36, 41, 112, 114, 118 and 129. Polynucleotides of the
CC invention are useful for treating diseases chosen from end stage liver
CC disease, hepatocellular carcinoma, anhidrotic ectoderm dysplasia,
CC degenerative neurological disorders, anaemia, ataxia, haemochromatosis,
CC sideroblastic anaemia and spinocerebellar ataxia. elf sequences are
CC useful for treating liver disorders chosen from cholestasis, biliary
CC stones, liver obstruction, stricture, primary biliary cirrhosis and
CC primary sclerosing cholangitis. praja-1 sequences are useful for
CC detecting colon cancer. The present sequence is mouse early liver
CC developmental protein.
XX
SQ Sequence 2154 AA;

Query Match 56.4%; Score 124; DB 8; Length 2154;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 SGSPSPATGVPPVNNOLBGRFLODERAVQKTTTKVNSHLAVSCRITDLYTDL 69
Db 10 SGSPSPATGVPPVNNOLBGRFLODERAVQKTTTKVNSHLAVSCRITDLYTDL 69
Db 70 RDGRMLIKLLEVLSEGERLPKPTKGRMRHCLLENVDKALQFLKEQVHLENNGSHDIVDGN 129
QY 130 HRLT 133
Db 70 RDGRMLIKLLEVLSEGERLPKPTKGRMRHCLLENVDKALQFLKEQVHLENNGSHDIVDGN 129
QY 130 HRLT 133
Db 130 HRLT 133

RESULT 8
ADZ85101
ID ADZ85101 standard; protein; 2154 AA.
XX
XX ADZ85101;
XX
XX 14-JUL-2005 (first entry)
XX
XX Full-length FHOS interacting protein, SEQ ID 150.
XX
XX FHOS; antiinflammatory; neuroprotective; cytosolic; cardiovascular-Gen.;
KM antidiabetic; inflammation; neurodegenerative disease; cancer;
KM cardiovascular disease; diabetes mellitus.
XX
XX Mus musculus.
OS
XX US2005100966-A1.
XX
XX 12-MAY-2005.
XX
XX 19-MAR-2004; 2004US-00805684.
XX
XX 02-APR-2003; 2003US-045936P.
PR

PR 02-APR-2003; 2003US-0460103P.
 PR 03-JUN-2003; 2003US-0455766P.
 XX
 XX (SAKA/) SAKAMOTO T.
 PA (TAKE/) TAKEDA S.
 XX
 PI Sakamoto T, Takeda S;
 XX WPI, 2005-345401/35.
 DR
 XX Novel isolated protein comprising FHO5 or its homolog, derivative or
 PT fragment, interacting with protein chosen from group of GROUP1 e.g.
 PT MRNP23, MRNP59 or MRBD7(627), useful for screening its modulator.
 XX
 PS Disclosure; SEQ ID NO 150; 163pp; English.
 XX
 XX The invention relates to a novel isolated protein comprising a first
 CC protein, which is FHO5 or its homolog, derivative or fragment,
 CC interacting with a second protein chosen from a group of GROUP1, e.g.
 CC MRNP23, MRNP59 or MRBD7(627), where the interaction is through a complex
 CC or covalent bond, or any other intermolecular interaction. The invention
 CC further comprises: a method for producing the FHO5-interacting protein; a
 CC method for detecting the FHO5-interacting protein in a sample; a method
 CC for determining whether a compound is capable of modulating an
 CC interaction between a first polypeptide (FHO5) or its homolog, derivative
 CC of fragments) and a second polypeptide as mentioned in the FHO5-
 CC interacting protein; and a method for modulating the function or activity
 CC of the FHO5-interacting protein in cells of a specific tissue of a
 CC mammal. The FHO5-interacting protein has the activities:
 CC antiinflammatory, neuroprotective, cytoskeletal, cardiovascular-Gen., and
 CC antidiabetic. The FHO5-interacting protein is useful for selecting its
 CC modulators. The method for modulating the function or activity of the
 CC FHO5-interacting protein in cells of a specific tissue is useful for
 CC treating inflammatory diseases, neurodegenerative diseases, cancer,
 CC cardiovascular diseases or diabetes mellitus. This sequence represents an
 CC FHO5 interacting protein for use in the novel protein complex of the
 CC invention.
 XX
 SQ Sequence 2154 AA;
 Query Match 56.4%; Score 124; DB 9; Length 2154;
 Best Local Similarity 100.0%; Pred. No. 1.3e-112;
 Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 SGPPLSAAYTGQVPYNNQLEGRPKQLODEREAVOKKFTKWNSHLARVSCRTTDLTYDL 69
 DB 10 SGPPLSAAYTGQVPYNNQLEGRPKQLODEREAVOKKFTKWNSHLARVSCRTTDLTYDL 69
 QY 70 RDGRMLIKLEVLVSGRLPKPTKGRMRHICLENVDKALQFLKEQRVHLENMGSMDIVDGN 129
 DB 70 RDGRMLIKLEVLVSGRLPKPTKGRMRHICLENVDKALQFLKEQRVHLENMGSMDIVDGN 129
 QY 130 HRLT 133
 DB 130 HRLT 133
 RESULT 9
 ID ABU70467 standard; protein; 321 AA.
 XX
 AC ABU70467;
 XX
 DT 10-JUN-2003 (first entry)
 XX
 DE Human adipocyte Selected Interacting domain, SID, #98.
 XX
 KM Human; prey; adipocyte; SID; selected interacting domain; anorectic;
 KM antidiabetic; protein-protein interaction; diabetes;
 KW yeast 2-hybrid assay; metabolic disorder; obesity.
 XX
 OS Homo sapiens.
 XX

PN W0200286122-A2.
 XX
 XX 31-OCT-2002.
 PD
 XX
 XX 14-MAR-2002; 2002WO-EP003768.
 PP
 XX 14-MAR-2001; 2001US-0275734P.
 PR
 XX (HYBR-) HYBRIGENICS.
 PA
 PI Legrain P, Davlet L;
 XX WPI, 2003-103412/09.
 DR
 DR N-PSDB; ACA57011.
 XX
 PT New complex between two interacting proteins in adipocyte cells, useful
 PT for identifying selected interacting domains that modulate protein
 PT interactions, or for preventing or treating metabolic disorders such as
 XX obesity or diabetes.
 XX
 PS Claim 6; Page 144; 382pp; English.
 XX
 XX The invention relates to a complex between two interacting proteins in
 CC adipocyte cells, given in the specification. The proteins are identified
 CC by selecting a bait protein from a known adipocyte marker and then
 CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
 CC members of an adipocyte cDNA library. The proteins are designated SID
 CC (RTM) (selected interacting domains) proteins. Also included are a
 CC polynucleotide encoding a polypeptide in the adipocyte cells, a
 CC recombinant host cell expressing at least one of the interacting
 CC polypeptides of the complex, selecting a modulating compound in adipocyte
 CC cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
 CC sequences given in the specification (including its fragment or variant),
 CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
 CC given in the specification (including its fragment or variant), a vector
 CC comprising the SID (RTM) polynucleotide, a recombinant host cell
 CC comprising the vector, a protein chip comprising the polypeptides and a
 CC record comprising all or part of the data, listed in the specification.
 CC The complex, polypeptides, polynucleotides and compounds are useful for
 CC preventing or treating metabolic disorders such as obesity or diabetes.
 CC The polynucleotides are useful as probes or primers. The complex is
 CC particularly useful for identifying selected interacting domains (SID
 CC (RTM)) for screening drugs that modulate the protein interaction, thus
 CC exhibiting the therapeutic effect. The present sequence represents a SID
 CC (prey) protein of the invention
 XX
 SQ Sequence 321 AA;
 Query Match 44.1%; Score 97; DB 6; Length 321;
 Best Local Similarity 100.0%; Pred. No. 1.1e-86;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 37 DEREAVOKKFTKWNSHLARVSCRTTDLTYDLRDGRMLIKLEVLVSGRLPKPTKGRMR 96
 DB 23 DEREAVOKKFTKWNSHLARVSCRTTDLTYDLRDGRMLIKLEVLVSGRLPKPTKGRMR 82
 QY 97 IHLENVDKALQFLKEQRVHLENMGSMDIVDGNHRLT 133
 DB 83 IHLENVDKALQFLKEQRVHLENMGSMDIVDGNHRLT 119
 RESULT 10
 ID ABU70517 standard; protein; 341 AA.
 XX
 AC ABU70517;
 XX
 DT 10-JUN-2003 (first entry)
 XX
 DE Human adipocyte Selected Interacting domain, SID, #148.
 XX
 KM Human; prey; adipocyte; SID; selected interacting domain; anorectic;
 KM antidiabetic; protein-protein interaction; diabetes;
 KW

KW Yeast 2-hybrid assay; metabolic disorder; obesity.
XX
OS Homo sapiens.
FN WO200286122-A2.
XX
PD 31-OCT-2002.
XX
PF 14-MAR-2002; 2002WO-EP003766.
XX
PR 14-MAR-2001; 2001US-0275734P.
XX
PA (HYBR-) HYBRIGENICS.
XX
PI Legrain P, Daviet L;
XX
DR WPI, 2003-103412/09.
XX
N-PSDB; ACAS7061.
XX
PT New complex between two interacting proteins in adipocyte cells, useful
PT for identifying selected interacting domains that modulate protein
PT interactions, or for preventing or treating metabolic disorders such as
PT obesity or diabetes.
XX
PS Claim 6; Page 156-157; 382pp; English.
XX
CC The invention relates to a complex between two interacting proteins in
CC adipocyte cells, given in the specification. The proteins are identified
CC by selecting a bait protein from a known adipocyte marker and then
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC members of an adipocyte cDNA library. The proteins are designated SID
CC (RTM) (selected interacting domains) proteins. Also included are a
CC polynucleotide encoding a polypeptide in the adipocyte cells, a
CC recombinant host cell expressing at least one of the interacting
CC polypeptides of the complex, selecting a modulating compound in adipocyte
CC cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
CC sequences given in the specification (including its fragment or variant),
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC given in the specification (including its fragment or variant), a vector
CC comprising the SID (RTM) polynucleotide, a recombinant host cell
CC comprising the vector, a protein chip comprising the polypeptides and a
CC record comprising all or part of the data, listed in the specification.
CC The complex, polypeptides, polynucleotides and compounds are useful for
CC preventing or treating metabolic disorders such as obesity or diabetes.
CC The polynucleotides are useful as probes or primers. The complex is
CC particularly useful for identifying selected interacting domains (SID
CC (RTM)) for screening drugs that modulate the protein interaction, thus
CC exhibiting the therapeutic effect. The present sequence represents a SID
CC (prey) protein of the invention
XX
SQ Sequence 341 AA;
XX
Query Match 44.1%; Score 97; DB 6; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.1e-86;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 37 DEREAIVOKKFTKWNVSHLARSCTITDYLDRGRMLIKLEVLSEGRLPKPTKGRMR 96
DB 39 DEREAIVOKKFTKWNVSHLARSCTITDYLDRGRMLIKLEVLSEGRLPKPTKGRMR 98
QY 97 IHCLENVDKALQFLKEQRVHLENNGSHDIYDGNHRLT 133
DB 99 IHCLENVDKALQFLKEQRVHLENNGSHDIYDGNHRLT 135
XX
RESULT 11
ID ADJ70625 standard; protein; 1630 AA.
XX
AC ADJ70625;
XX
DT 06-MAY-2004 (first entry)
XX

DE Human heat mitochondrial protein as a therapeutic target SegID2431.
XX
KW Mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cyostatic.
XX
OS Homo sapiens.
XX
PN WO2003087768-A2.
XX
PD 23-OCT-2003.
XX
PF 04-APR-2003; 2003WO-US010870.
XX
PR 12-APR-2002; 2002US-0372843P.
XX
PR 17-JUN-2002; 2002US-0389987P.
XX
PR 20-SEP-2002; 2002US-0412418P.
XX
PA (MITO-) MITOKOR.
XX
PA (BUCK-) BUCK INSR AGE RES.
XX
PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW;
PI Warnock DE;
XX
DR WPI; 2003-845369/78.
XX
PT Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX
PS Claim 1; SEQ ID NO 2431; 180pp; English.
XX
CC This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nootropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cyostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
SQ Sequence 1630 AA;
XX
Query Match 44.1%; Score 97; DB 7; Length 1630;
Best Local Similarity 100.0%; Pred. No. 4.7e-86;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 37 DEREAIVOKKFTKWNVSHLARSCTITDYLDRGRMLIKLEVLSEGRLPKPTKGRMR 96
DB 50 DEREAIVOKKFTKWNVSHLARSCTITDYLDRGRMLIKLEVLSEGRLPKPTKGRMR 109
QY 97 IHCLENVDKALQFLKEQRVHLENNGSHDIYDGNHRLT 133
DB 110 IHCLENVDKALQFLKEQRVHLENNGSHDIYDGNHRLT 146
XX
RESULT 12
ID ABP65068 standard; protein; 2364 AA.
XX
AC ABP65068;
XX

Db 110 IHCLENVDKALQFLKEQRVHLENMGSHDIVDGNHRLT 146

RESULT 14

ADZ70337 standard; protein; 2364 AA.

XX ADZ70337;

XX 30-JUN-2005 (first entry)

XX Human protein from lung cancer marker gene SPTEN1.

XX Tumor marker; lung tumor; cytostatic; neoplasm; expression;

XX DNA microarray.

XX Homo sapiens.

XX WO2005032495-A2.

XX 14-APR-2005.

XX 01-OCT-2004; 2004WO-US034163.

XX 03-OCT-2003; 2003US-0508355P.

XX (FARB) BAYER PHARM CORP.

XX Taylor I, Pauloski NR, Bigwood D;

XX WPI; 2005-285325/29.

XX N-PSDB; ADZ70336.

XX Providing a patient diagnosis for lung cancer comprises comparing the level of expression of genes or gene products in a biological sample from the patient with that from a normal individual.

XX Claim 3; SEQ ID NO 22; 60pp; English.

XX The invention relates to providing a patient diagnosis for lung cancer comprising comparing the level of expression of genes or gene products in a biological sample from the patient with the level of expression of genes or gene products in a biological sample from a normal individual.

XX Also included are distinguishing between normal and disease tissues, monitoring the response of a patient being treated for lung cancer by administering an anti-cancer agent, identifying a compound useful for the treatment of lung cancer and an array for distinguishing between normal and disease tissues (comprising 2 or more probes corresponding to 2 or more genes selected from any of the 200 nucleotide sequences given in the specification, or 2 or more polypeptides comprising any of the 200 amino acid sequences given in the specification). In providing a patient diagnosis for lung cancer, one or more genes are selected from any of the 200 nucleotide sequences as mentioned in the specification, or one or more gene products are polypeptides selected from any of the 20 amino acid sequences mentioned in the specification. The methods are useful for detecting and treating lung cancer. These may also be used for designing, identifying and optimizing therapeutics for cancer. The present sequence represents a protein from one of the 200 lung cancer marker genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2364 AA;

Query Match 44.1%; Score 97; DB 9; Length 2364;

Best Local Similarity 100.0%; Pred. No. 5.8e-86;

Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 DERAVVQKTKTKVNSHARVSCRITLTYDIDBGRMLITLLEVLSESRPKPKTKGMR 96

DB 50 DERAVVQKTKTKVNSHARVSCRITLTYDIDBGRMLITLLEVLSESRPKPKTKGMR 109

OY 97 IHCLENVDKALQFLKEQRVHLENMGSHDIVDGNHRLT 133

Db 110 IHCLENVDKALQFLKEQRVHLENMGSHDIVDGNHRLT 146

RESULT 15

AAW81642 standard; protein; 1120 AA.

XX AAW81642;

XX 01-MAR-1999 (first entry)

XX Mouse elf protein.

XX Elf; mouse; liver development; signal transduction; biliary tree;

XX cholestasis; biliary stone; hepatic obstruction; stricture;

XX primary biliary cirrhosis; primary sclerosing cholangitis; therapy.

XX Mus sp.

XX WO9848827-A1.

XX 05-NOV-1998.

XX 30-APR-1998; 98WO-US008656.

XX 30-APR-1997; 97US-00841349.

XX (MISH/) MISHRA L.

XX Mishra L;

XX WPI; 1999-009382/01.

XX N-PSDB; AAW84424.

XX New isolated early liver development genes - used to develop products for treating, e.g. liver disease, hepatocellular carcinoma, degenerative neurological disorders, anaemia, ataxia or haemochromatosis.

XX Example 2; Fig 16a; 92pp; English.

XX This is the amino acid sequence of an elf protein of the foetal murine liver, as deduced from an isolated partial cDNA clone (see AAW64424). The invention provides early developing liver proteins (see AAW81638-42) and the genes coding for them (see AAW64410-24). The genes were isolated from embryonic cDNA libraries using subtractive hybridisation. 3 Elf genes (see AAW64411-13) were identified. The elf proteins (see also AAW81639-40) are probably important for the formation of the biliary tree during early liver development. They are used in a claimed method to treat cholestasis, biliary stones, liver obstruction, stricture, primary biliary cirrhosis and primary sclerosing cholangitis. Early developing liver proteins and nucleic acids can also be used in the diagnosis of liver diseases and other disorders, including those relating to oncogenesis and tissue repair.

XX Sequence 1120 AA;

Query Match 40.9%; Score 90; DB 2; Length 1120;

Best Local Similarity 100.0%; Pred. No. 2.4e-79;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 RLTTLELLEVRROOEERERKRRPPSPPNKTVSEASQOMDSKGNQVONGIPRAGGS 190

DB 1031 RLTTLELLEVRROOEERERKRRPPSPPNKTVSEASQOMDSKGNQVONGIPRAGGS 1090

QY 191 PRVSYRSQTYONYKNFNSRRRTASDHSWGM 220

DB 1091 PRVSYRSQTYONYKNFNSRRRTASDHSWGM 1120

RESULT 16

ADJ70287 standard; protein; 2106 AA.

XX ADJ70287;
 AC
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human heat mitochondrial protein as a therapeutic target SegID2093.
 XX
 KW mitochondrial; human; screening assay; diabetes mellitus;
 KW Huntington's disease; osteoarthritis;
 KW Leber's hereditary optic neuropathy; LHON;
 KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
 KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
 KW neuroprotective; nootropic; anticonvulsant; antiarthritic;
 KW osteopathic; ophthalmological; cyostatic.
 XX
 OS Homo sapiens.
 XX
 PN WO2003087768-A2.
 XX
 PD 23-OCT-2003.
 XX
 PF 04-APR-2003; 2003WO-US010870.
 XX
 PR 12-APR-2002; 2002US-0372843P.
 PR 17-JUN-2002; 2002US-038987P.
 PR 20-SEP-2002; 2002US-0412418P.
 XX
 PA (MITO-) MITOKOR.
 PA (BUCK-) BUCK INST AGE RES.
 PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
 PI Warnock DE;
 PI WPI; 2003-845369/78.
 DR
 XX Identifying a mitochondrial target for drug screening assays and for
 PT treating diseases associated with altered mitochondrial function,
 PT comprises detecting a modified polypeptide in a sample and correlating
 PT with the disease.
 PS
 PS Claim 1: SEQ ID NO 2093; 180pp; English.
 XX
 CC This invention relates to novel mitochondrial targets that can be used
 CC for therapeutic intervention in treating a disease associated with
 CC altered mitochondrial function. Specifically, it refers to a method for
 CC identifying proteins of the human heart mitochondria proteome that are
 CC useful for drug screening assays, as well as therapeutic targets. The
 CC present invention describes a method for identifying such proteins that
 CC can be used in the treatment of various diseases associated with altered
 CC mitochondrial function including diabetes mellitus, Huntington's disease,
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
 CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
 CC compositions have neuroprotective, nootropic, antidiabetic,
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
 CC cyostatic activities. This polypeptide sequence is a human heart
 CC mitochondrial protein of the invention.
 CC
 SQ Sequence 2106 AA;
 XX
 Query Match 17.3%; Score 38; DB 7; Length 2106;
 Best Local Similarity 100.0%; Pred. No. 7.4e-28;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 MRHICLENVDKALQFLKEORVHLNMGSHDIVDGNHRL 132
 DB 108 MRHICLENVDKALQFLKEORVHLNMGSHDIVDGNHRL 145
 XX
 RESULT 17
 ADX06608 standard; protein; 2137 AA.
 XX

AC ADX06608;
 XX
 XX 21-APR-2005 (first entry)
 DT
 XX
 DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 1173.
 XX
 KW cyostatic; cyclin-dependent kinase; cdk; biomarker.
 XX
 OS Homo sapiens.
 XX
 PN WO2005012875-A2.
 XX
 PD 10-FEB-2005.
 XX
 PF 29-JUL-2004; 2004WO-US024424.
 XX
 PR 29-JUL-2003; 2003US-0490890P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
 PI P-PSDB; ADX06607.
 DR WPI; 2005-163068/17.
 DR P-PSDB; ADX06607.
 XX
 PT Biomarkers useful for predicting or determining the response of a mammal
 PT to a cancer treatment comprising administration of a modulator of cyclin-
 PT dependent kinase activity.
 PS
 PS Claim 5; SEQ ID NO 1173; 141pp; English.
 XX
 CC This invention describes a novel method of predicting or determining
 CC whether a mammal will respond or is responding to an anti-cancer agent
 CC that modulates cyclin-dependent kinase (cdk) activity. The method
 CC comprises measuring the level of one or more biomarkers selected from
 CC 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
 CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
 CC invention is utilized in a kit for determining or predicting whether
 CC patient would be susceptible or resistant to treatment by an agent
 CC modulating cdk activity. The invention also describes a method for
 CC utilizing individualized genetic profiles for treating diseases and
 CC disorders based on patient's response and molecular level, specialized
 CC microarrays comprising the biomarkers described, antibodies directed
 CC against the biomarkers and a cell culture model to identify biomarkers.
 CC The cdk modulator is preferably N-5-[(5-(1,1-dimethylethyl)-2-
 CC oxazolyl)methyl]-thio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
 CC tartaric acid salt. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
 CC sequence represents a biomarker used in the method of the invention.
 CC
 SQ Sequence 2137 AA;
 XX
 Query Match 17.3%; Score 38; DB 9; Length 2137;
 Best Local Similarity 100.0%; Pred. No. 7.4e-28;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 MRHICLENVDKALQFLKEORVHLNMGSHDIVDGNHRL 132
 DB 108 MRHICLENVDKALQFLKEORVHLNMGSHDIVDGNHRL 145
 XX
 RESULT 18
 ABR41636
 ID ABR41636 standard; protein; 2141 AA.
 AC ABR41636;
 XX
 DT 02-JUN-2003 (first entry)
 XX
 DE Human DITRP cytoskeletal protein.
 XX
 KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;

KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging;
 KW cytoskeletal protein.
 KW
 OS Homo sapiens.
 XX
 PN WO200297031-A2.
 XX
 PD 05-DEC-2002.
 XX
 PF 27-MAR-2002; 2002WO-US010056.
 XX
 PR 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.
 PR 29-MAR-2001; 2001US-0280068P.
 PR 16-MAY-2001; 2001US-0291280P.
 PR 17-MAY-2001; 2001US-0291829P.
 PR 17-MAY-2001; 2001US-0291843P.
 PR 19-JUN-2001; 2001US-0299428P.
 PR 20-JUN-2001; 2001US-0299776P.
 PR 20-JUN-2001; 2001US-0300001P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;
 PI Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Ameshey SR;
 PI Daugherty SC, Dem TC, Nguyen DA, Kleeefeld Y, Gerstin EH;
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;
 PI Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;
 XX
 DR WPI; 2003-129518/12.
 DR N-PSDB; ACC46573.
 XX
 PT Novel human diagnostic and therapeutic polypeptide useful for identifying
 PT test compound which specifically binds to a polypeptide encoded by human
 PT diagnostic and therapeutic polynucleotide, and to induce antibodies.
 XX
 PS .Claim 27; SEQ ID NO 1171; 591pp; English.
 XX
 CC The invention relates to novel human diagnostic and therapeutic
 CC polynucleotides designated ditrp (ACC46080-ACC46749) and to their encoded
 CC proteins (DITRP; ABR41136-ABR41812). The invention also relates to
 CC polynucleotide sequences at least 90% identical to the ditrp CDNA
 CC sequences of the invention; recombinant vectors, host cells and
 CC transgenic organisms comprising a ditrp nucleic acid sequence; the
 CC recombinant production of DITRP proteins; antibodies specific for DITRP
 CC proteins; microarrays comprising ditrp nucleic acid sequences; methods of
 CC detecting ditrp nucleotide and protein sequences; methods of screening
 CC for compounds which specifically bind a DITRP protein; and methods of
 CC assessing the toxicity of test compounds using a ditrp hybridisation
 CC probe. Ditrp nucleic acid sequences and DITRP proteins may be used in the
 CC diagnosis of a wide variety of conditions including cancer and other cell
 CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,
 CC viral, fungal or parasitic infections; hormonal disorders; metabolic
 CC disorders; neurological disorders; gastrointestinal disorders; transport
 CC disorders; and connective tissue disorders. They may also be used to
 CC screen for modulators of protein activity or gene expression. DITRP
 CC proteins can additionally be used in analysis of the proteome of a tissue
 CC or cell type and to induce antibodies. The ditrp nucleic acids are
 CC additionally useful in somatic or germline gene therapy of the disorders
 CC mentioned above, as a source of antisense sequences, as a source of
 CC probes and primers, in genotyping and identification of individuals, in
 CC the generation of transgenic animal models of human disease or knock in
 CC humanised animals, in toxicological testing, and in transcript imaging.
 CC The present sequence represents a DITRP protein which is a cytoskeletal
 CC protein. Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 2141 AA;
 Query Match 17.3%; Score 38; DB 6; Length 2141;
 Best Local Similarity 100.0%; Pred. NO. 7.5e-28;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 MRHCLENVKALQFLKEQVHLENMGSHDIVDGNHRL 132
 |||||
 DB 112 MRHCLENVKALQFLKEQVHLENMGSHDIVDGNHRL 149
 |||||
 RESULT 19
 AAU01183
 ID AAU01183 standard; protein; 2387 AA.
 XX
 AC AAU01183;
 XX
 DT 07-SEP-2001 (first entry)
 XX
 DE Rat glutamate transporter associated protein GTRAP4-41.
 XX
 KW Rat; glutamate transporter associated protein; GTRAP4-41;
 KW glutamate transporter protein; chloride transport; neurodegeneration;
 KW cytoskeletal stability; nervous system disorder; schizophrenia;
 KW spinocerebellar ataxia type 1; SCAL; GABA metabolism.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 2048..2049
 FT /note="Encoded by GTAGGAGC"
 XX
 PN WO200130968-A2.
 XX
 PD 03-MAY-2001.
 XX
 PF 23-OCT-2000; 2000WO-US029431.
 XX
 PR 23-OCT-1999; 99US-0161007P.
 PR 22-MAY-2000; 2000US-0206157P.
 XX
 PA (UNIV) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
 XX
 PI Rochstein JD, Jackson M, Lin G, Law R, Orlov I;
 PI WPI; 2001-300498/31.
 DR N-PSDB; AAS04273.
 XX
 PT Novel substantially pure glutamate transporter associated polypeptide
 PT which modulates intracellular glutamate transport, interacts with
 PT glutamate transporter protein and has expression pattern in brain.
 XX
 PS Claim 7; Fig 15; 116pp; English.
 XX
 CC The present sequence representing novel rat glutamate transporter
 CC associated protein GTRAP4-41 modulates intracellular glutamate transport,
 CC interacts with a glutamate transporter protein and has an expression
 CC pattern in the brain. The GTRAP proteins are also involved in mediating
 CC chloride transport and cytoskeletal stability. The invention also
 CC provides methods for identifying a compound that modulates a cellular
 CC response mediated by a GTRAP protein and a compound that inhibits the
 CC interaction of a glutamate transporter protein and a GTRAP protein. A
 CC compound which modulates the activity of a glutamate transporter
 CC associated polypeptide or interaction with a glutamate transporter
 CC protein is useful for treating a disorder associated with glutamate
 CC transport or chloride flux, preferably a nervous system disorder such as
 CC neurodegeneration, spinocerebellar ataxia type 1 (SCAL), schizophrenia,
 CC epilepsy or a disorder of GABA metabolism. Such compound could be an
 CC antisense oligonucleotide (AAS04278) which can be used for modulating
 CC glutamate transport in a subject
 XX
 SQ Sequence 2387 AA;

Query Match 12.3%; Score 27; DB 4; Length 2387;
Best Local Similarity 100.0%; Pred. No. 5.7e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LPKPTKGRMRHICLENVDKALQFLKEQ 113
DB 103 LPKPTKGRMRHICLENVDKALQFLKEQ 129

RESULT 20

AAV05494
ID AAV05494 standard; protein; 2390 AA.

AC AAV05494;

DT 08-JUL-1999 (first entry)

DE BetaIII spectrin protein sequence.

BetaIII spectrin; intracellular sequestration modulation; SNAIS;

spectrin-ankyrin-adaptor protein trafficking/anchoring system;

integral membrane protein; secretory protein; endoplasmic reticulum;

cis-Golgi apparatus; medial-Golgi apparatus; trans-Golgi apparatus;

plasma cell membrane; cystic fibrosis.

Mammalia.

WO9916875-A1.

08-APR-1999.

30-SEP-1998; 98WO-US020364.

30-SEP-1997; 97US-0060559P.

(UYVA) UNIV YALE.

Morrow JS, Devareajan P;

WPI; 1999-277085/23.

N-PSDB; AAX36583.

Modulating intracellular sequestration of a selected integral membrane or

secretory protein into a transport vehicle.

Claim 21; Fig 19e; 129pp; English.

This sequence represents a mammalian betaIII spectrin. The invention relates to a method for modulating intracellular sequestration of a selected integral membrane or secretory protein into a transport vehicle (for transport from the endoplasmic reticulum to the cis-Golgi apparatus, or from the cis-Golgi to the medial-Golgi apparatus, or from the medial-Golgi to the trans-Golgi apparatus or from the trans-Golgi apparatus to the plasma cell membrane) using the spectrin-ankyrin-adaptor protein trafficking/anchoring system (SNAIS). Compounds that inhibit or enhance intracellular sequestration of selected integral membrane or secretory protein into a transport vehicle are useful for ameliorating effects of disease caused by a relative excess or deficit in the amount of the particular protein. Identification of e.g. the binding domain for CFTF (cystic fibrosis transmembrane conductance regulator) responsible for cystic fibrosis, in the spectrin-ankyrin-adaptor protein trafficking/anchoring system (SNAIS), by the methods would allow early analysis of possible transport enhancing reagents that might the clinical severity of this disease. The method does not attempt to suppress the synthesis of a given protein, only its delivery to the correct cellular tissue compartment. Intracellular expression of RNA is not required and the method lends itself to high-throughput in vitro screening assays

Sequence 2390 AA;

Query Match 12.3%; Score 27; DB 2; Length 2390;

Best Local Similarity 100.0%; Pred. No. 5.7e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LPKPTKGRMRHICLENVDKALQFLKEQ 113
DB 103 LPKPTKGRMRHICLENVDKALQFLKEQ 129

RESULT 21

ABP41709
ID ABP41709 standard; protein; 806 AA.

AC ABP41709;

DT 22-AUG-2002 (first entry)

DE Human ovarian antigen HBODF41, SEQ ID NO:2841.

Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

ovarian cancer; breast cancer; tumour; reproductive system disorder;

infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;

inflammatory condition; immune disorder; blood disorder;

cardiovascular disorder; respiratory disorder; neurological disorder;

gastrointestinal disorder; urinary system disorder; drug screening;

gene therapy; chromosome mapping; forensic analysis;

antibody preparation; cytostatic; immunomodulatory; neuroprotective;

antiinflammatory; gynaecological; reproductive.

Homo sapiens.

WO200200677-A1.

03-JAN-2002.

07-JUN-2001; 2001WO-US018569.

07-JUN-2000; 2000US-0209467P.

(HUMA-) HUMAN GENOMB SCT INC.

Blaze CE, Rosen CA;

WPI; 2002-147878/19.

N-PSDB; ABQ54786.

Isolated nucleic acid molecules encoding novel ovarian polypeptides,

useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian

cancer), immune disorders, cardiovascular disorders and neurological

diseases.

Claim 11; SEQ ID NO 2841; 2922pp; English.

The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosis or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the

CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 806 AA;

Query Match 10.5%; Score 23; DB 5; Length 806;
 Best Local Similarity 100.0%; Pred. No. 2e-13;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 RLTLLELRVROQEEERKRP 153
 DB 699 RLTLLELRVROQEEERKRP 721

RESULT 22

ABB61876 ID ABB61876 standard; protein; 2291 AA.

XX ABB61876;

AC 26-MAR-2002 (first entry)

DT Drosophila melanogaster polypeptide SEQ ID NO 12420.

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL05979.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions.

XX Disclosure; SEQ ID NO 12420; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL6176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL6175) and the encoded proteins (ABBS7737-

XX ABB12072). The sequence data for this patent did not form part of the

XX printed specification, but was obtained in electronic format directly

XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 2291 AA;

XX Query Match 8.6%; Score 19; DB 4; Length 2291;

XX Best Local Similarity 100.0%; Pred. No. 4.2e-09;

XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 LKLEVLGGERLPKPTKG 93
 DB 84 LKLEVLGGERLPKPTKG 102

RESULT 23
 ID ADF74518
 ADF74518 standard; peptide; 15 AA.

XX ADF74518;

XX 26-FEB-2004 (first entry)

DE Peptide from the C-terminus of the murine elf 3 protein.

XX murine; mouse; early liver development; end stage liver disease; elf 1-3;

XX 11yor-1 (145); pk; protein 106; praja-1; hepatocyte lineage; cholestasis;

XX biliary stone; hepatic obstruction; stricture; primary biliary cirrhosis;

XX primary sclerosing cholangitis; gene therapy;

XX anidrotic ectoderm dysplasia; hepatocellular carcinoma; anemia; ataxia;

XX neurological disorder; haemochromatosis; hepatotropic; cytostatic;

XX neuroprotective; antihaemic; cardiant; antibody.

XX Mus musculus.

XX US6642362-B1.

XX 04-NOV-2003.

XX 01-NOV-1999; 99US-00431184.

XX 30-APR-1997; 97US-00841349.

XX 30-APR-1998; 98WO-US008656.

XX (MISH/) MISHRA L.

XX Mishra L;

XX WPI; 2003-851362/79.

XX This invention relates to novel genes and encoded proteins thereof.

XX isolated during early liver development that are useful in the diagnosis

XX and treatment of end stage liver disease and other disorders.

XX Specifically, it refers to genes that encode proteins such as elf 1-3,

XX 11yor-1 (145), pk, protein 106 and praja-1. The present invention

XX describes the characterisation of these early liver development proteins,

XX and also methods to raise peptide specific antibodies that are useful as

XX markers, as well as for tracing hepatocyte lineage. Furthermore, elf

XX proteins 1-3 are useful in treating disorders including cholestasis,

XX biliary stones, hepatic obstruction, stricture, primary biliary cirrhosis

XX or primary sclerosing cholangitis. In addition, through using gene

XX therapy, praja-1, 11yor-1 (145) and pk can be used to treat anidrotic

XX ectoderm dysplasia, hepatocellular carcinoma and other diseases including

XX haemochromatosis. Accordingly, these proteins can be described as

XX hepatotropic, cytostatic, neuroprotective, antihaemic and cardiant. This

XX peptide sequence is from the C-terminus of the murine elf 3 protein, used

XX to raise rabbit anti-mouse antibodies of the invention. NOTE: This

XX sequence differs from the murine Elf-3 peptide sequence given in the

XX sequence listing (SeqID 20).

XX Sequence 15 AA;

XX Query Match 6.8%; Score 15; DB 7; Length 15;

XX Best Local Similarity 100.0%; Pred. No. 4.9e-07;

XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      206 FNSRRTASDHSWGM 220
          |||||
Db      1 FNSRRTASDHSWGM 15
```

RESULT 24

ADQ88396
ID ADQ88396 standard; peptide; 15 AA.

AC ADQ88396;

DT 07-OCT-2004 (first entry)

DE Mouse e1f-3 epitope peptide #2.

KM Early liver developmental protein; liver disease;
KM hepatocellular carcinoma; anhidrotic ectoderm dysplasia;
KM degenerative neurological disorder; anaemia; ataxia; haemochromatosis;
KM adrenolethal anaemia; spinocerebellar ataxia; cholestasis;
KM biliary stone; liver obstruction; stricture; primary biliary cirrhosis;
KM primary sclerosing cholangitis; colon cancer; therapy; mouse; elf;
KM embryonic liver fodrin; beta spectrin; gliptoe.

OS Mus musculus.

PN US2004142354-A1.

PD 22-JUL-2004

PF 30-OCT-2003; 2003US-00695994.

PR 30-APR-1997; 97US-00841349.

PR 01-NOV-1999; 99US-00431184.

PA (MISH/) MISHRA L.

PI M18hra L;

WPI; 2004-552558/53.

PT Novel isolated early

PT disease, hepatocellular carcinoma, anemia and ataxia.

PS Example 2; Page 17; 85pp; English.

CC The invention provides novel early liver developmental proteins such as

CC protein 106 and praja-1 and nuclear

includ:genes 20, 36, 41, 112, 114, 118 and 129. Polynucleotides of the invention are useful for treating diseases chosen from end stage liver disease, hepatocellular carcinoma, anidrotic ectoderm dysplasia, degenerative neurological disorders, anaemia, ataxia, haemochromatosis, sideroblastic anaemia and spinocerebellar ataxia. elf sequences are useful for treating liver disorders chosen from cholestasis, biliary stones, liver obstruction, stricture, primary biliary cirrhosis and primary sclerosing cholangitis. praja.1 sequences are useful for detecting colon cancer. The present sequence is mouse elf -3 epitope peptide.

SQ Sequence 15 AA;

Query Match	6.8%	Score 15;	DB 8;	Length 15;
Best Local Similarity	100.0%;	Pred. NO.	4.9e-07;	
Matches 15;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY 206 FNSRRTASDHSWSGM 220

Db 1 FNSRRTASDHSWSGM 15

RESULT 25
ADV13561
ID ADV13561 standard; peptide; 17 AA.

AC ADV13561;

DT 10-MAR-2005 (first entry)

DE Human phosphorylated peptide from phosphoprotein #1499

KM Alzheimer's disease; neuroprotective; nootropic; degeneration; tumor;
KM neoplasm; neurological disease; phosphorylation; protein sequencing;
KM phosphoprotein.
...

OS Homo sapiens.

PN WO2004108948-A2

PD 16-DEC-2004.

PF 04-JUN-2004; 2004WO-US017613.

PR 04-JUN-2003; 2003US-0476010P.

PA (HARD) HARVARD COLLEGE.

PI Gygi SP,

DR WPI; 2005-031720/03

PT Characterizing phosphorylated polypeptides in a sample
PT digesting the polypeptides with a protease thus generating test peptides
PT and collecting a fraction of test peptides that enriched for positively charged peptides.

PS Claim 16; Page 87; 123pp; English

The invention relates to characterizing phosphorylated polypeptides in a sample comprising digesting the polypeptides with a protease thus generating test peptides, and collecting a fraction of test peptides that enriched for positively charged peptides. Also included are a method (comprising determining the presence, absence or level of one or more phosphorylated peptides as identified above in cells having a cell state and determining the degree of correlation between the presence, absence or level of phosphorylated polypeptide with the cell state), an isolated peptide of 5-50 amino acids comprising an amino acid sequence that is a subsequence of any of the protein sequences given in the specification (and which comprise a phosphorylation site within the subsequence), an isolated polypeptide selected from any of the polypeptides listed in the specification and is modified at a modification site, an isolated peptide comprising a mass spectral peak signature. a method for identifying a treatment that modulates phosphorylation of an amino acid in a target polypeptide, a method for generating a peptide standard, a pair of peptide standards comprising the peptide obtained (where the peptide is phosphorylated and a corresponding peptide comprising an identical amino acid sequence but which is not phosphorylated), a system (comprising a computer memory comprising data files storing information relating to the identifying characteristics of positively charged peptides, and a data analysis module capable of executing instructions for organizing and/or searching the data files), a computer program product (comprising data relating to the identifying characteristics of positively charged peptides and comprising instructions for organizing and/or searching the data), and a method for identifying N-terminal peptides in a sample. The method is useful for characterizing phosphorylated polypeptides in a sample. The present sequence is a peptide from a human phosphoprotein, containing a phosphorylation site, identified by the method of the invention.

SQ Sequence 17 AA;

Query Match	6.8%;	Score 15;	DB 9;	length 17;
Best Local Similarity	100.0%;	Pred. No. 5.4e-07;		

	Matches	15;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	178 QVSONGLPAEQGSPR 192 									
Db	3 QVSONGLPAEQGSPR 17									

RESULT 26

ID	ADVA13562 standard; peptide; 17 AA.
XX AC	ADVA13562;
XX XX	10-MAR-2005 (first entry)
DE XX	Human phosphorylated peptide from phosphoprotein #1500.
KM KM	Alzheimers disease; neuroprotective; nocitropic; degeneration; tumor; neoplasm; neurological disease; phosphorylation; protein sequencing; phosphoprotein.
OS XX	Homo sapiens.
PN XX	WO2004108948-A2.
PD XX	16-DEC-2004.
PF XX	04-JUN-2004; 2004WO-US017613.
PR XX	04-JUN-2003; 2003US-0476010P.
PA XX	(HARD) HARVARD COLLEGE.
PI XX	Gygi SP;
DR XX	WPI; 2005-031720/03.

Characterizing phosphorylated polypeptides in a sample comprises digesting the polypeptides with a protease thus generating test peptides, and collecting a fraction of test peptides that enriched for positively charged peptides.

Claim 16; Page 87; 123pp; English.

The invention relates to characterizing phosphorylated polypeptides in a sample comprising digesting the polypeptides with a protease thus generating test peptides, and collecting a fraction of test peptides that enriched for positively charged peptides. Also included are a method comprising determining the presence, absence or level of one ore more phosphorylated peptides as identified above in cells having a cell state and determining the degree of correlation between the presence, absence or level of phosphorylated polypeptide with the cell state), an isolated peptide of 5-50 amino acids comprising an amino acid sequence that is a subsequence of any of the protein sequences given in the specification (and which comprise a phosphorylation site within the subsequence)', an isolated polypeptide selected from any of the polypeptides listed in the specification and is modified at a modification site, an isolated peptide comprising a mass spectral peak signatures. a method for identifying a treatment that modulates phosphorylation of an amino acid in a target polypeptide, a method for generating a peptide standard, a pair of peptide standards comprising the peptide obtained (where the peptide is phosphorylated and a corresponding peptide comprising an identical amino acid sequence but which is not phosphorylated)', a system (comprising a computer memory comprising data files storing information relating to the identifying characteristics of positively charged peptides, and a data analysis module capable of executing instructions for organizing and/or searching the data files), a computer program product (comprising data relating to the identifying characteristics of positively charged peptides) and comprising instructions for organizing and/or searching the data), and a method for identifying N-terminal peptides in a sample. The method is useful for characterizing phosphorylated polypeptides in a sample. The present sequence is a peptide from a human phosphoprotein, containing a phosphorylation site identified by the method of the

CC invention.
XX
SQ Sequence 17 AA;
6.8%; Score 15; DB 9; Length 17;
Query Match
Best Local Similarity 100.0%; Pred. No. 5.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 178 QVSONGLPAEQGSPP 192
DB 3 QVSONGLPAEQGSPP 17

RESULT 27
AAW81644
ID AAW81644 standard; peptide; 14 AA.
XX
XX AAW81644;
AC
DT 01-MAR-1999 (first entry)
DE
XX
XX Mouse elf-1 C-terminal peptide used to raise antibody.
XX
XX Elf-1; mouse; liver development; signal transduction; biliary tree;
KM cholestasis; biliary stone; hepatic obstruction; stricture;
KM primary biliary cirrhosis; primary sclerosing cholangitis; therapy;
KM antibody.
XX
XX Mus sp.
OS
XX WO9848827-A1.
XX
XX 05-NOV-1998.
XX
XX 30-APR-1998; 98WO-US008656.
XX PF
XX 30-APR-1997; 97US-00841349.
XX FR
XX (MISH/) MISHRA L.
XX PA
XX
XX Mishra L;
XX
XX WPI; 1999-009382/01.
XX DR
XX
XX New isolated early liver development genes - used to develop products for
XX treating, e.g. liver disease, hepatocellular carcinoma, degenerative
XX neurological disorders, anaemia, ataxia or haemochromatosis.
XX
XX Claim 29; Page 25; 92pp; English.

CC This peptide corresponds to amino acids 2140-2154 of the mouse elf-1
XX protein (see AAW81638). Claimed peptides (see AAW81643-52) derived from
CC early liver development proteins (see AAW81638-41) of the invention have
CC been used to raise antibodies. It is an object of the invention to
CC provide proteins which are characteristic of early liver development, and
CC to raise antibodies from these proteins which will be useful as markers,
CC and will be useful in methods of identifying such proteins and peptides,
CC tracing hepatocyte lineage, and treating liver disease. Elf-1 is used in
CC a claimed method to treat cholestasis, biliary stones, liver obstruction,
CC stricture, primary biliary cirrhosis and primary sclerosing cholangitis
XX
XX
SQ Sequence 14 AA;

Query Match 6.4%; Score 14; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY 206 FNSRTASDHSMSG 219
DB 1 FNSRTASDHSMSG 14

RESULT 28

ADP74507
ID ADF74507 standard; peptide; 14 AA.
XX
AC ADF74507;
XX
XX 26-FEB-2004 (first entry)
DE Peptide from the C-terminus of the murine elf 3 protein (SeqID20).
XX
XX murine; mouse; early liver development; end stage liver disease; elf 1-3;
KM liyor-1 (145); pk; protein 106; praja-1; hepatocyte lineage; cholestasis;
KM biliary stone; hepatic obstruction; stricture; primary biliary cirrhosis;
KM primary sclerosing cholangitis; gene therapy;
KM anhidrotic ectoderm dysplasia; hepatocellular carcinoma; anaemia; ataxia;
KM neurological disorder; haemochromatosis; hepatotropic; cytostatic;
KM neuroprotective; anti-anemic; cardiac; antibody.
XX
OS Mus musculus.
XX
PN US6642362-B1.
XX
PD 04-NOV-2003.
XX
PF 01-NOV-1999; 99US-00431184.
XX
PR 30-APR-1997; 97US-00841349.
PR 30-APR-1998; 98WO-US008656.
XX
PA (MISH/) MISHRA L.
XX
PI Mishra L;
PI
DR WPI; 2003-851362/79.
XX
XX
PT New antibodies recognizing early liver development proteins, useful as
PT markers, in identifying peptides and proteins having early liver
PT development characteristics, tracing hepatocyte lineage or treating liver
PT disease.
XX
XX
PS Claim 1; SEQ ID NO 20; 82pp; English.
XX
XX This invention relates to novel genes and encoded proteins thereof,
CC isolated during early liver development that are useful in the diagnosis
CC and treatment of end stage liver disease and other disorders.
CC Specifically, it refers to genes that encode proteins such as elf 1-3,
CC liyor-1 (145), pk, protein 106 and praja-1. The present invention
CC describes the characterisation of these early liver development proteins,
CC and also methods to raise peptide specific antibodies that are useful as
CC markers, as well as for tracing hepatocyte lineage. Furthermore, elf
CC proteins 1-3 are useful in treating disorders including cholestasis,
CC biliary stones, hepatic obstruction, stricture, primary biliary cirrhosis
CC or primary sclerosing cholangitis. In addition, through using gene
CC therapy, praja-1, liyor-1 (145) and pk can be used to treat anhidrotic
CC ectoderm dysplasia, hepatocellular carcinoma and other diseases including
CC anaemia, ataxia, degenerative neurological disorders and
CC haemochromatosis. Accordingly, these proteins can be described as
CC hepatotropic, cytostatic, neuroprotective, anti-anemic and cardiac. This
CC peptide sequence is from the C-terminus of the murine elf 3 protein, used
CC to raise rabbit anti-mouse antibodies of the invention. NOTE: This
CC sequence differs from the murine Elf-3 peptide sequence given in example
CC 2.
XX
SQ Sequence 14 AA;
XX
Query Match 6.4%; Score 14; DB 7; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 29
ADQ88379
ID ADQ88379 standard; peptide; 14 AA.
XX
AC ADQ88379;
XX
XX 07-OCT-2004 (first entry)
DE Mouse elf epitope peptide #2.
XX
XX Early liver developmental protein; liver disease;
KM hepatocellular carcinoma; anhidrotic ectoderm dysplasia;
KM degenerative neurological disorder; anaemia; ataxia; haemochromatosis;
KM sideroblastic anaemia; spinocerebellar ataxia; cholestasis;
KM biliary stone; liver obstruction; stricture; primary biliary cirrhosis;
KM primary sclerosing cholangitis; colon cancer; therapy; mouse; elf;
KM embryonic liver fodrin; beta spectrin; epitope.
XX
XX Mus musculus.
XX
PN US2004142354-A1.
XX
PD 22-OCT-2003; 2003US-00695994.
XX
PF 30-OCT-2003; 2003US-00695994.
XX
PR 30-APR-1997; 97US-00841349.
PR 30-APR-1998; 98WO-US008656.
PR 01-NOV-1999; 99US-00431184.
XX
PA (MISH/) MISHRA L.
XX
PI Mishra L;
PI
DR WPI; 2004-552558/53.
XX
XX
PT Novel isolated early liver developmental proteins such as elf, liyor-1
PT (145), pk, protein 106 and praja-1, useful for treating end stage liver
PT disease, hepatocellular carcinoma, anemia and ataxia.
XX
XX
PS Claim 17; SEQ ID NO 20; 85pp; English.
XX
XX The invention provides novel early liver developmental proteins such as
CC elf (embryonic liver fodrin or beta spectrin) 1-3, liyor-1 (145), pk,
CC protein 106 and praja-1 and nucleic acid molecules encoding such
CC proteins. Other genes coding for early liver developmental proteins
CC include genes 20, 36, 41, 112, 114, 118 and 129. Polynucleotides of the
CC invention are useful for treating diseases chosen from end stage liver
CC disease, hepatocellular carcinoma, anhidrotic ectoderm dysplasia,
CC degenerative neurological disorders, anaemia, ataxia, haemochromatosis,
CC sideroblastic anaemia and spinocerebellar ataxia. elf sequences are
CC useful for treating liver disorders chosen from cholestasis, biliary
CC stones, liver obstruction, stricture, primary biliary cirrhosis and
CC primary sclerosing cholangitis. praja-1 sequences are useful for
CC detecting colon cancer. The present sequence is mouse elf epitope
CC peptide.
XX
SQ Sequence 14 AA;
XX
Query Match 6.4%; Score 14; DB 8; Length 14;
Best Local Similarity 100.0%; Pred. No. 4.4e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 30
ADY65644
ID ADY65644 standard; protein; 597 AA.
XX
AC ADY65644;

XX 02-JUN-2005 (first entry)
DT
XX
XX S. mansoni protein SEQ ID 1062.
DE
XX
XX Schistosomacide; schistosomiasis; psychiatric disorder; vaccine;
KM diagnosis.
XX
XX Schistosoma mansoni.
OS
XX
XX WO2005023979-A2.
PN
XX
XX 17-MAR-2005.
PD
XX
XX 10-SEP-2004; 2004WO-BR000170.
PF
XX
XX 11-SEP-2003; 2003US-0502277P.
PR
XX
XX (AMPA-) FUNDACAO AMPARO A PESQUISA DO ESTADO.
PA
XX
XX Verjovski-Almeida S, Leite LCC, Farias LP, Miyasato PA, Kawano T;
PI Demarco R, Garcia JCL, Martins EAL, Ho PU, Nascimento ALTO; V;
PI Dias-Neto B, Secubal JC, Menck CFM, Madeira AMB, Rodrigues V;
PI Gargioni C;
XX
XX WPI; 2005-223357/23.
DR
XX
XX New isolated nucleic acid molecule encoding a Schistosoma mansoni
PT protein, useful for as a vaccine or for preventing, diagnosing, or
PT treating Schistosoma mansoni infection.
PT
XX
XX Disclosure; SEQ ID NO 1062; 52bp; English.
PS
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
CC Schistosoma mansoni protein, or its portion which is at least 20 amino
CC acids in length. Also included are an expression vector comprising the
CC nucleic acid operably linked to a promoter, a recombinant cell
CC transformed/transfected with the nucleic acid (or expression vector), an
CC immunogenic composition comprising the nucleic acid in combination with a
CC pharmaceutical adjuvant or carrier, an isolated S. mansoni protein
CC comprising an amino acid sequence encoded by the nucleic acid, an
CC immunogenic composition comprising the isolated S. mansoni protein, an
CC isolated antibody which specifically binds to the isolated S. mansoni
CC protein, a hybridoma cell line which produces the isolated antibody, a
CC method for determining a S. mansoni infection in a subject and a computer
CC readable medium having recorded in it a nucleic acid molecule from S.
CC mansoni genome. The nucleic acid molecule is useful for preventing,
CC diagnosing, or treating S. mansoni infection (schistosomiasis). It can
CC also be used as vaccine against S. mansoni. The present sequence is a
CC Schistosoma mansoni protein of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. Also, SEQ ID 2141-2152 are
CC mentioned in the specification but are not included in the sequence
CC listing.
XX
XX Sequence 597 AA;
SQ

Query Match 6.4%; Score 14; DB 9; Length 597;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VQKKTFTKWNVSHL 55
DB 64 VQKKTFTKWNVSHL 77

RESULT 31
AAB50652
ID AAB50652 standard; protein; 2326 AA.
XX
XX AAB50652;
AC
XX

DT 19-MAR-2001 (first entry)
XX
XX C. elegans spectrin beta-chain/fodrin protein SEQ ID NO:12.
DE
XX
XX Human; Caenorhabditis elegans; UNC-5; splice variant; nematode worm;
KM protein-protein interaction; identification.
XX
XX Caenorhabditis elegans.
OS
XX
XX WO2000073328-A2.
PN
XX
XX 07-DEC-2000.
PD
XX
XX 02-JUN-2000; 2000WO-BP005108.
PF
XX
XX 01-JUN-1999; 99GB-00012755.
PR
XX
XX (DEVG-) DEVGEN NV.
PA
XX
XX Van Criekeinge W, Roelens I, Bogaert T, Vervaeke P;
PI WPI; 2001-016508/02.
PI
XX
XX N-PSDB; AAC90918.
DR
XX
XX Three variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) and a
PT human unc-5HS1 cDNA, useful in yeast two hybrid experiments for
PT identifying unknown human cDNAs which encode proteins that interact with
PT the human unc-5C protein.
PT
XX
XX Claim 37; Page 110-118; 246bp; English.
PS
XX
XX The present invention describes 3 variants of human unc-5C cDNAs (unc-
CC 5Cb, unc-5Cc and unc-5C8) which correspond to alternatively spliced unc-
CC 5C transcripts, and a human unc-5HS1 cDNA which shares homology with the
CC Rattus norvegicus unc-5HS1 cDNA. Also described are assays based on
CC protein-protein-interactions between the unc-5 protein and a variety of
CC different interacting proteins. The unc-5C variant cDNAs and unc-5HS1
CC cDNA are useful in methods for identifying compounds which reduce or
CC inhibit the lethal phenotype associated with the expression of the unc-5
CC death domain in yeast. They are also useful in yeast two hybrid
CC experiments for identifying unknown human cDNAs which encode proteins
CC that interact with the human unc-5C protein. AAC90914 to AAC90971 and
CC AAB50646 to AAB50693 represent sequences used in the exemplification of
CC the present invention.
XX
XX Sequence 2326 AA;
SQ

Query Match 6.4%; Score 14; DB 4; Length 2326;
Best Local Similarity 100.0%; Pred. No. 0.00026;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VQKKTFTKWNVSHL 55
DB 110 VQKKTFTKWNVSHL 123

RESULT 32
ADN23721
ID ADN23721 standard; protein; 2326 AA.
XX
XX ADN23721;
AC
XX
XX 02-DEC-2004 (first entry)
DT
XX
XX Bacterial polypeptide #6374.
DE
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KM cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KM patchogen tolerance; pest tolerance; plant disease resistance;
KM cell cycle pathway modification; plant growth regulator;
KM homologous recombination; seed oil yield; protein yield; carbohydrate;
KM nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KM bacterial polypeptide.

XX Bacteria.
OS US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 6374; 122pp; English.
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polypeptide used in the
XX scope of the invention. Note: The sequence data for this patent did not
XX form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
SQ Sequence 2326 AA;
Query Match 6.4%; Score 14; DB 8; Length 2326;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 VQKKTFTKWNVSHL 55
DB 110 VQKKTFTKWNVSHL 123
RESULT 33
ID AEA55025 standard; protein; 385 AA.
XX AEA55025;
AC AEA55025;
XX 11-AUG-2005 (first entry)
DT Human actin-cross-linking protein, ACF7, protein fragment, SEQ ID NO: 1.
XX

KW Plasma membrane; diagnosis; therapeutic; cancer; cytostatic; neoplasm;
KM actin-cross-linking protein.
XX Homo sapiens.
XX WO2005052182-A2.
XX 09-JUN-2005.
XX 25-NOV-2004; 2004WO-1L001085.
XX 26-NOV-2003; 2003US-0524885P.
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX Linial M, Inberg A, Bledi Y;
XX WPI; 2005-418017/42.
XX PIR; S66292.
XX Characterizing proteins present in a plasma membrane of a cell, useful in
XX identifying diagnostic markers and potential drugs, comprises subjecting
XX a cell to a protease treatment.
XX Claim 25; SEQ ID NO 1; 196pp; English.
XX The present invention relates to a method of characterizing proteins
XX present in the plasma membrane (PM) of live cells. The proteins of the
XX invention are useful in identifying diagnostic markers and potential
XX drugs. The invention is useful for identifying drugs for diagnosing and
XX treating disorders such as cancer which are associated with abnormal
XX representation of cell surface proteins. The present sequence is human
XX actin-cross-linking protein, ACF7, protein fragment.
SQ Sequence 385 AA;
Query Match 5.0%; Score 11; DB 9; Length 385;
Best Local Similarity 100.0%; Pred. No. 0.069;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 VQKKTFTKWNV 52
DB 96 VQKKTFTKWNV 106
RESULT 34
ID AAY59242 standard; protein; 897 AA.
XX AAY59242;
AC AAY59242;
XX 31-MAR-2000 (first entry)
DT A rod shortened dystrophin (deltaDysH4).
XX Muscular dystrophy; rod domain; adeno-associated virus; AAV;
XX dystrophin gene; truncated.
XX Homo sapiens.
XX JP1318467-A.
XX 24-NOV-1999.
XX 08-MAY-1998; 98JP-00142134.
XX 08-MAY-1998; 98JP-00142134.
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX (KOKU-) KOKURITSU SEISHIN SHINKAI CENT SOCHO.
XX WPI; 2000-100771/09.
XX N-PSDB; AA248571.
DR

XX A gene encoding a shortened dystrophin - useful for the treatment of
PT muscular dystrophy.

XX Disclousre; Page 32-33; 44pp; Japanese.

CC The invention provides a gene for the treatment of muscular dystrophy
CC having at least one rod repeat structure of hinge 1, hinge 4 and rod
CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene
CC and a gene-introducing medium consisting of an adeno-associated virus
CC (AAV) vector or lentivirus vector containing the rod shortened dystrophin
CC genes can be used for the genetic treatment of muscular dystrophy of low
CC immune reaction. The present sequence represents a rod shortened
CC dystrophin

XX Sequence 897 AA;

Query Match 5.0%; Score 11; DB 3; Length 897;

Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 42 VOKTFTKWN 52
|||
16 VOKTFTKWN 26

Db 16 VOKTFTKWN 26

RESULT 35

AAVS9241
ID AAVS9241 standard; protein; 900 AA.

XX AAVS9241;

XX 31-MAR-2000 (first entry)

XX A rod shortened dystrophin (deltadysh1).

XX Muscular dystrophy; rod domain; adeno-associated virus; AAV;

XX dystrophin gene; truncated.

XX Homo sapiens.

XX JP11318467-A.

XX 24-NOV-1999.

XX 08-MAY-1998; 98JP-00142134.

XX 08-MAY-1998; 98JP-00142134.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX (KOKU-) KOKURITSU SEISHIN SHINKAI CENT SOCHO.

XX WPI; 2000-100771/09.

XX N-PSDB; AAZ48570.

XX A gene encoding a shortened dystrophin - useful for the treatment of
PT muscular dystrophy.

XX Disclousre; Page 26-28; 44pp; Japanese.

CC The invention provides a gene for the treatment of muscular dystrophy
CC having at least one rod repeat structure of hinge 1, hinge 4 and rod
CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene
CC and a gene-introducing medium consisting of an adeno-associated virus
CC (AAV) vector or lentivirus vector containing the rod shortened dystrophin
CC genes can be used for the genetic treatment of muscular dystrophy of low
CC immune reaction. The present sequence represents a rod shortened
CC dystrophin

XX Sequence 900 AA;

Query Match 5.0%; Score 11; DB 3; Length 900;

Best Local Similarity 100.0%; Pred. No. 0.14;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 42 VOKTFTKWN 52
|||
16 VOKTFTKWN 26

Db 16 VOKTFTKWN 26

RESULT 36

AAVS9237
ID AAVS9237 standard; protein; 1092 AA.

XX AAVS9237;

XX 31-MAR-2000 (first entry)

XX A rod shortened dystrophin (deltadyshx2).

XX Muscular dystrophy; rod domain; adeno-associated virus; AAV;

XX dystrophin gene; truncated.

XX Homo sapiens.

XX JP11318467-A.

XX 24-NOV-1999.

XX 08-MAY-1998; 98JP-00142134.

XX 08-MAY-1998; 98JP-00142134.

XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX (KOKU-) KOKURITSU SEISHIN SHINKAI CENT SOCHO.

XX WPI; 2000-100771/09.

XX N-PSDB; AAZ48566.

XX A gene encoding a shortened dystrophin - useful for the treatment of
PT muscular dystrophy.

XX Claim 8; Page 14-16; 44pp; Japanese.

CC The invention provides a gene for the treatment of muscular dystrophy
CC having at least one rod repeat structure of hinge 1, hinge 4 and rod
CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene
CC and a gene-introducing medium consisting of an adeno-associated virus
CC (AAV) vector or lentivirus vector containing the rod shortened dystrophin
CC genes can be used for the genetic treatment of muscular dystrophy of low
CC immune reaction. The present sequence represents a rod shortened
CC dystrophin

XX Sequence 1092 AA;

Query Match 5.0%; Score 11; DB 3; Length 1092;

Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 42 VOKTFTKWN 52
|||
16 VOKTFTKWN 26

Db 16 VOKTFTKWN 26

RESULT 37

AAVS9240
ID AAVS9240 standard; protein; 1201 AA.

XX AAVS9240;

XX 31-MAR-2000 (first entry)

XX A rod shortened dystrophin (deltadyshx3).

XX Muscular dystrophy; rod domain; adeno-associated virus; AAV;

XX dystrophin gene; truncated.

```

XX OS Homo sapiens.
XX XX JP11318467-A.
XX PN 24-NOV-1999.
XX PD 08-MAY-1998; 98JP-00142134.
XX PF 08-MAY-1998; 98JP-00142134.
XX PR 08-MAY-1998; 98JP-00142134.
XX PA (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.
XX PA (KOKU-) KOKURITSU SEISHIN SHINKAI CENT SOCHO.
XX DR MPI; 2000-100771/09.
XX DR N-PSDB; AA248569.
XX PT A gene encoding a shortened dystrophin - useful for the treatment of
XX PT muscular dystrophy.
XX PS Disclosure; Page 26-28; 44pp; Japanese.
XX CC The invention provides a gene for the treatment of muscular dystrophy
XX CC having at least one rod repeat structure of hinge 1, hinge 4 and rod
XX CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene
XX CC and a gene-introducing medium consisting of an adeno-associated virus
XX CC (AAV) vector or lentivirus vector containing the rod shortened dystrophin
XX CC genes can be used for the genetic treatment of muscular dystrophy of low
XX CC immune reaction. The present sequence represents a rod shortened
XX CC dystrophin
XX SQ Sequence 1201 AA;

Query Match 5.0%; Score 11; DB 3; Length 1201;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 VOKKTFTKMN 52
DB 16 VOKKTFTKMN 26

RESULT 38
AAV59238
ID AAV59238 standard; protein; 1310 AA.
XX AC AAV59238;
XX DT 31-MAR-2000 (first entry)
XX DE A rod shortened dystrophin (deltadySAx11).
XX KW Muscular dystrophy; rod domain; adeno-associated virus; AAV;
XX KM dystrophin gene; truncated.
XX OS Homo sapiens.
XX XX JP11318467-A.
XX PN 24-NOV-1999.
XX PD 08-MAY-1998; 98JP-00142134.
XX PF 08-MAY-1998; 98JP-00142134.
XX PR 08-MAY-1998; 98JP-00142134.
XX PA (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.
XX PA (KOKU-) KOKURITSU SEISHIN SHINKAI CENT SOCHO.
XX DR MPI; 2000-100771/09.
XX DR N-PSDB; AA248567.
XX PT A gene encoding a shortened dystrophin - useful for the treatment of
XX PT muscular dystrophy.

```

```

XX XX Claim 8; Page 18-21; 44pp; Japanese.
XX CC The invention provides a gene for the treatment of muscular dystrophy
XX CC having at least one rod repeat structure of hinge 1, hinge 4 and rod
XX CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene
XX CC and a gene-introducing medium consisting of an adeno-associated virus
XX CC (AAV) vector or lentivirus vector containing the rod shortened dystrophin
XX CC genes can be used for the genetic treatment of muscular dystrophy of low
XX CC immune reaction. The present sequence represents a rod shortened
XX CC dystrophin
XX SQ Sequence 1310 AA;

Query Match 5.0%; Score 11; DB 3; Length 1310;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 VOKKTFTKMN 52
DB 16 VOKKTFTKMN 26

RESULT 39
AAV59239
ID AAV59239 standard; protein; 1310 AA.
XX AC AAV59239;
XX DT 31-MAR-2000 (first entry)
XX DE A rod shortened dystrophin (deltadySAH3).
XX KW Muscular dystrophy; rod domain; adeno-associated virus; AAV;
XX KM dystrophin gene; truncated.
XX OS Homo sapiens.
XX XX JP11318467-A.
XX PN 24-NOV-1999.
XX PD 08-MAY-1998; 98JP-00142134.
XX PF 08-MAY-1998; 98JP-00142134.
XX PR 08-MAY-1998; 98JP-00142134.
XX PA (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.
XX PA (KOKU-) KOKURITSU SEISHIN SHINKAI CENT SOCHO.
XX DR MPI; 2000-100771/09.
XX DR N-PSDB; AA248568.
XX PT A gene encoding a shortened dystrophin - useful for the treatment of
XX PT muscular dystrophy.
XX OS Claim 8; Page 22-24; 44pp; Japanese.
XX XX The invention provides a gene for the treatment of muscular dystrophy
XX CC having at least one rod repeat structure of hinge 1, hinge 4 and rod
XX CC domain of dystrophin gene and having a base sequence of 4.5 kb. The gene
XX CC and a gene-introducing medium consisting of an adeno-associated virus
XX CC (AAV) vector or lentivirus vector containing the rod shortened dystrophin
XX CC genes can be used for the genetic treatment of muscular dystrophy of low
XX CC immune reaction. The present sequence represents a rod shortened
XX CC dystrophin
XX SQ Sequence 1310 AA;

Query Match 5.0%; Score 11; DB 3; Length 1310;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 VOKKTFTKMN 52

```


Db 16 VQKFTKVN 26

|||||

RESULT 40

ID ADN22598 standard; protein; 1355 AA.

AC ADN22598;

DT 02-DEC-2004 (first entry)

DE Bacterial polypeptide #5251.

XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

OS Bacteria.

PN US2003233675-A1.

PD 18-DEC-2003.

PF 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

PA (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

PI WPI, 2004-061375/06.

PT New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 5251, 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

SQL Sequence 1355 AA;

Query Match 5.0%; Score 11; DB 8; Length 1355;

Best Local Similarity 100.0%; Pred. No. 0.2;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 42 VQKFTKVN 52

79 VQKFTKVN 89

RESULT 41

ID ADN22598 standard; protein; 1355 AA.

AC ADN22598;

DT 02-DEC-2004 (first entry)

DE Bacterial polypeptide #5252.

XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.

OS Bacteria.

PN US2003233675-A1.

PD 18-DEC-2003.

PF 20-FEB-2003; 2003US-00369493.

PR 21-FEB-2002; 2002US-0360039P.

PA (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

PI WPI, 2004-061375/06.

PT New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.

PS Claim 1; SEQ ID NO 5252, 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

CC Sequence 1355 AA;

Query Match 5.0%; Score 11; DB 8; Length 1355;
Best Local Similarity 100.0%; Pred. No. 0.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 VQKKTFTKWN 52
|||
79 VQKKTFTKWN 89

RESULT 42

AEA24046
ID AEA24046 standard; protein; 2000 AA.

XX AEA24046;

XX 11-AUG-2005 (first entry)

XX Human PRO polypeptide SEQ ID NO 588.

XX Immune disorder; PRO; Antiinflammatory; Dermatological;

XX Immunosuppressive; Antirheumatic; Antiarthritic; Osteopathic;

XX Muscular-Gen.; Vasotropic; Antianemic; Antipsoriatic; Immunostimulant.

XX Homo sapiens.

XX WO2005051988-A2.

XX 09-JUN-2005.

XX 02-MAR-2004; 2004WO-US006460.

XX 03-MAR-2003; 2003US-0451884P.

XX (GENTH) GENENTECH INC.

XX Abbas A, Bodary S, Clark H, Schoenfeld J, Williams PM, Wood WI;

XX Wu TD;

XX WPI; 2005-417958/42.

XX N-PSDB; AEA24045.

XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and

XX treating an immune related disorder, e.g. systemic lupus erythematosus,

XX rheumatoid arthritis, osteoarthritis, autoimmune hemolytic anemia, or

XX psoriasis.

XX Disclosure; SEQ ID NO 588; 966pp; English.

XX The invention relates to an isolated nucleic acid. The polypeptide,

XX compound or composition, and methods are useful for diagnosing and

XX treating an immune related disorder, e.g. systemic lupus erythematosus,

XX rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis,

XX spondyloarthropathies, systemic sclerosis, idiopathic inflammatory

XX myopathies, Sjogren's syndrome, systemic vasculitis, sarcoidosis,

XX autoimmune hemolytic anemia, autoimmune or immune-mediated skin diseases

XX including bullous skin diseases, erythema multiforme and contact

XX dermatitis, psoriasis, lymphadenopathy, splenomegaly and leukopenia. The

XX present sequence represents the amino acid sequence of a human PRO

XX polypeptide.

XX Sequence 2000 AA;

XX Query Match 5.0%; Score 11; DB 9; Length 2000;

XX Best Local Similarity 100.0%; Pred. No. 0.29;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 42 VQKKTFTKWN 52
|||
70 VQKKTFTKWN 80

RESULT 43

AAP90373
ID AAP90373 standard; protein; 3685 AA.

XX AAP90373;

XX 29-MAR-1992 (first entry)

XX Sequence encoded by human muscular dystrophy (MD) cDNA.

XX Dystrophin; muscular dystrophy; probe; antibody; diagnosis; prenatal;

XX heterozygote; gene therapy; genetic screening; foetal screening.

XX Homo sapiens.

XX WO8906286-A.

XX 13-JUL-1989.

XX 16-DEC-1988; 88WO-US004504.

XX 22-DEC-1987; 87US-0036618.

XX (CHIL-) CHILDRENS MED CENT.

XX Kunkel LM, Monaco A, Hoffman EP, Koenig M;

XX WPI; 1989-220587/30.

XX N-PSDB; AAP90338.

XX Muscular dystrophy gene - used for prepn. of probes, dystrophic

XX polypeptide and antibodies for diagnosis and therapy of muscular

XX dystrophy.

XX Disclosure; Fig 8; 68pp; English.

XX The inventors claim an MD probe comprising a purified as NA SQ which

XX hybridises to at least a part of the MD gene; pure dystrophin (DS)

XX polypeptide, purified NA encoding DS and antibodies (Ab) to DS. The

XX probes are equal to or greater than 10b of one of 12 cDNA sequences

XX deposited as ATCC 58666-57677. The MD gene is human, or a murine Dmd gene

XX Sequence 3685 AA;

XX Query Match

XX Best Local Similarity 100.0%; Pred. No. 0.48;

XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX OY 42 VQKKTFTKWN 52
|||
16 VQKKTFTKWN 26

RESULT 44

AAP90290
ID AAP90290 standard; protein; 3685 AA.

XX AAP90290;

XX 25-MAR-2003 (revised)

XX 10-JAN-1990 (first entry)

XX Human Duchenne muscular dystrophy gene.

XX Duchenne muscular dystrophy; protein deletion; antiserum.

OS Homo sapiens.
XX
XX EP31514-A.
XX
XX PD 06-SEP-1989.
XX
XX PD 03-MAR-1989; 89EP-00302145.
XX
XX PF 04-MAR-1988; 88JP-00051313.
XX
XX PR 21-SEP-1988; 88JP-00237200.
XX
XX PA (AJIN) AJINOMOTO KK.
XX
XX PI Ishiguro T, Eguichi C;
XX
XX DR WPI, 1989-257828/36.
XX
XX PT Detecting human Duchenne muscular dystrophy - by assaying for the
XX associated protein deletion or defect using antibody to the protein.
XX
XX PS Claim 3; Page 15-16; 20pp; English.
XX
XX CC Peptides may be prepared consisting of all or part of the sequence and
XX used to produce Abs for detecting protein deletions or defects in the
XX CC gene. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-
XX CC 2003 to correct PA field.)
XX
XX SQ Sequence 3685 AA;

Query Match 5.0%; Score 11; DB 1; Length 3685;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTKWN 52
DB 16 VOKKFTKWN 26

RESULT 45
ADL83110
ID ADL83110 standard; protein; 3685 AA.
XX
XX AC ADL83110;
XX
XX DT 17-JUN-2004 (first entry)
XX
XX DE Human PRO84704, SEQ ID 312.
XX
XX KW Immunosuppressive; Cytostatic; Antiarthritic; Antirheumatic; Antianemic;
XX Antiallergic; Muscular; Neuroprotective; Nephrotoxic; Antiinflammatory;
XX KW Gene therapy; PRO; B cell related disorder; cancer;
XX KW Immune-mediated inflammatory disease; human.
XX
XX OS Homo sapiens.
XX
XX PN WO2004024097-A2.
XX
XX PD 25-MAR-2004.
XX
XX PF 15-SEP-2003; 2003WO-US029097.
XX
XX PR 16-SEP-2002; 2002US-0411392P.
XX
XX PA (GERTH) GENENTECH INC.
XX
XX PI Chiu H, Clark H, Dennis K, Fong S, Schoenfeld JR, Wood WI;
XX PI Wu TD;
XX
XX DR WPI, 2004-329389/30.
XX
XX DR N-PSDB; ADL83109.
XX
XX PT New PRO polypeptide, useful for diagnosing and treating a B cell related
XX disorder, e.g. Burkitt's lymphoma, rheumatoid arthritis, autoimmune

PT mediated hemolytic anemia, myasthenia gravis or ankylosing spondylitis.
XX
XX PS Claim 10; Fig 312; 695pp; English.
XX
XX CC The present invention relates to PRO proteins and their coding sequences.
XX CC The PRO proteins are useful for diagnosing and treating a B cell related
XX CC disorder, e.g. X-linked infantile hypogammaglobulinemia, polysaccharide
XX CC antigen unresponsiveness, selective IGA deficiency, selective IGM
XX CC deficiency, selective deficiency of Igg subclasses, immunodeficiency with
XX CC hyper IgM, transient hypogammaglobulinemia of infancy, Burkitt's
XX CC lymphoma, intermediate lymphoma, follicular lymphoma, type II
XX CC hypersensitivity, rheumatoid arthritis, autoimmune mediated haemolytic
XX CC anaemia, myasthenia gravis, hypoadrenocorticism, glomerulonephritis, or
XX CC ankylosing spondylitis. The PRO proteins are also useful for preparing a
XX CC medicament for treating a condition that is responsive to the PRO
XX CC protein, e.g. cancer or immune-mediated inflammatory diseases. The PRO
XX CC coding sequences are useful as hybridization probes in chromosome and
XX CC gene mapping, in preparing PRO proteins, or in generating transgenic
XX CC animals or knockout animals, which in turn are useful in the development
XX CC and screening of therapeutically useful reagents.
XX
XX SQ Sequence 3685 AA;

Query Match 5.0%; Score 11; DB 8; Length 3685;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTKWN 52
DB 16 VOKKFTKWN 26

RESULT 46
ADN04004
ID ADN04004 standard; protein; 3685 AA.
XX
XX AC ADN04004;
XX
XX DT 01-JUL-2004 (first entry)
XX
XX DE Antipsoriatic protein sequence #197.
XX
XX KW antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
XX OS Homo sapiens.
XX
XX PN WO2004028479-A2.
XX
XX PD 08-APR-2004.
XX
XX PF 25-SEP-2003; 2003WO-US030907.
XX
XX PR 25-SEP-2002; 2002US-0414006P.
XX
XX PA (GERTH) GENENTECH INC.
XX
XX PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
XX PI Wu TD;
XX
XX DR WPI, 2004-305105/28.
XX
XX DR N-PSDB; ADN04003.
XX
XX PT New PRO nucleic acid or polypeptide, useful for preparing a
XX PT pharmaceutical composition for diagnosing or treating psoriasis in a
XX PT mammal.
XX
XX PS Claim 9; SEQ ID NO 398; 3069pp; English.
XX
XX CC The invention relates to novel polynucleotide and polypeptides for
XX CC treating psoriasis or a sequence having at least 80% identity to the
XX CC above sequences. The nucleic acid is useful for preparing a composition
XX CC for diagnosing or treating psoriasis in a mammal. This sequence
XX CC corresponds to one of the polypeptides of the invention.

XX Sequence 3685 AA;
SQ
Query Match 5.0%; Score 11; DB 8; Length 3685;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 VQKTFKTKWN 52
16 VQKTFKTKWN 26
Db
RESULT 47
AAU14603
ID AAU14603 standard; protein; 5373 AA.
XX
XX AAU14603;
AC
XX
XX 24-OCT-2001 (first entry)
DT
XX
XX Novel bone marrow polypeptide #2.
DE
XX Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
KM haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
KM wound healing; nutritional supplement; immune disorder;
KM severe combined immunodeficiency; SCID.
XX
XX Homo sapiens.
OS
XX WO200157187-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 05-FEB-2001; 2001WO-US003782.
PF
XX
XX 03-FEB-2000; 2000US-00496914.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 30-NOV-2000; 2000US-0250683P.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Ford JF, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;
PI Ren F, Drmanac RJ;
PI
XX WPI: 2001-488875/53.
DR N-PSDB; AAS22908.
XX
XX Nucleic acids encoding bone marrow polypeptides, useful in diagnostic and
PT gene therapy.
PT
XX
XX Claim 10; Page 198-208; 392pp; English.
PS
XX AAU14603-AAU14794 represent novel bone marrow polypeptides of the
CC invention. The proteins and corresponding coding sequences may be used in
CC the prevention, diagnosis and treatment of diseases associated with
CC inappropriate bone marrow polypeptide expression. For example, to treat
CC disorders associated with decreased expression by rectifying mutations or
CC deletions in a patient's genome that affect the activity of the
CC polypeptides by expressing inactive proteins or to supplement the
CC patient's own production of the polypeptide. Additionally, the nucleic
CC acids may be used to produce the polypeptide, by inserting the nucleic
CC acids into a host cell and culturing the cell to express the protein. The
CC nucleic acid and its complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and therefore which patients
CC may be in need of restorative therapy. The proteins may also be used as
CC antigens in the production of antibodies against bone marrow proteins and
CC in assays to identify modulators of their expression and activity. The
CC anti-bone marrow protein antibodies and antagonists may also be used to
CC down regulate expression and activity. The antibodies may also be used as
CC diagnostic agents for detecting the presence of the protein in samples
CC (e.g. by enzyme linked immunosorbant assay (ELISA)). The proteins may be

CC used to regulate haematopoiesis activity, and consequently in the
CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
CC such as wound healing; as a nutritional supplement; and in treatment of
CC immune disorders such as severe combined immunodeficiency (SCID)
XX
XX Sequence 5373 AA;
SQ
Query Match 5.0%; Score 11; DB 4; Length 5373;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 VQKTFKTKWN 52
44 VQKTFKTKWN 54
Db
RESULT 48
ADJ68935
ID ADJ68935 standard; protein; 5373 AA.
XX
XX ADJ68935;
AC
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX Human heat mitochondrial protein as a therapeutic target SegID741.
DE
XX
XX mitochondrial; human; screening assay; diabetes mellitus;
XX Huntington's disease; osteoarthritis;
XX Leber's hereditary optic neuropathy; LHON;
XX mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
XX myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
XX neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
XX osteopathic; ophthalmological; cyostatic.
XX
XX Homo sapiens.
OS
XX WO2003087768-A2.
PN
XX
XX 23-OCT-2003.
PD
XX
XX 04-APR-2003; 2003WO-US010870.
PF
XX
XX 12-APR-2002; 2002US-0372843P.
PR 17-JUN-2002; 2002US-0389987P.
PR 20-SEP-2002; 2002US-0412418P.
XX
XX (MITO-) MITOKOR.
PA (BUCK-) BUCK INST AGE RES.
PA
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GW,
PI Wamock DE;
PI
XX WPI: 2003-845369/78.
DR
XX
XX Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
PT with the disease.
XX
XX Claim 1; SEQ ID NO 741; 180pp; English.
PS
XX This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nootropic, antidiabetic,

CC anti-conulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.

XX Sequence 5373 AA;

Query Match 5.0%; Score 11; DB 7; Length 5373;

Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YY 42 VOKKFTKWN 52
|||
44 VOKKFTKWN 54

DB 42 VOKKFTKWN 52
|||
44 VOKKFTKWN 54

RESULT 49

ID AD285102 standard; protein; 5430 AA.

AC AD285102;

DT 14-JUL-2005 (first entry)

DE Full-length FOS interacting protein, SEQ ID 151.

XX FOS; antiinflammatory; neuroprotective; cytosolic; cardiovascular-Gen.;
XX antidiabetic; inflammation; neurodegenerative disease; cancer;

KM cardiovascular disease; diabetes mellitus.

XX Homo sapiens.

PN US2005100966-A1.

PD 12-MAY-2005.

PF 19-MAR-2004; 2004US-00805684.

PR 02-APR-2003; 2003US-0459936P.

PR 02-APR-2003; 2003US-0460103P.

PR 03-JUN-2003; 2003US-0455766P.

XX (SAKA/) SAKAMOTO T.

PA (TAKE/) TAKEDA S.

PI Sakamoto T, Takeda S;

DR WPI; 2005-345401/35.

PT Novel isolated protein comprising FOS or its homolog, derivative or
PT fragment, interacting with proteins chosen from group of GROUP1 e.g.
PT MRNP23, MRNP5 or MRNP7(627), useful for screening its modulator.

PS Disclosure; SEQ ID NO 151; 163pp; English.

XX The invention relates to a novel isolated protein comprising a first
XX protein, which is FOS or its homolog, derivative or fragment,
XX interacting with a second protein chosen from a group of GROUP1, e.g.

CC MRNP23, MRNP5 or MRNP7(627), where the interaction is through a complex
CC or covalent bond, or any other intermolecular interaction. The invention
CC further comprises: a method for producing the FOS-interacting protein; a
CC method for detecting the FOS-interacting protein in a sample; a method
CC for determining whether a compound is capable of modulating an
CC interaction between a first polypeptide (FOS) or its homolog, derivative
CC of fragments) and a second polypeptide as mentioned in the FOS-
CC interacting protein; and a method for modulating the function or activity
CC of the FOS-interacting protein in cells of a specific tissue of a
CC mammal. The FOS-interacting protein has the activities:

CC antiinflammatory, neuroprotective, cytosolic, cardiovascular-Gen., and
CC antidiabetic. The FOS-interacting protein is useful for selecting its
CC modulators. The method for modulating the function or activity of the
CC FOS-interacting protein in cells of a specific tissue is useful for
CC treating inflammatory diseases, neurodegenerative diseases, cancer.

CC cardiovascular diseases or diabetes mellitus. This sequence represents an

CC FOS interacting protein for use in the novel protein complex of the
CC invention.

XX Sequence 5430 AA;

Query Match 5.0%; Score 11; DB 9; Length 5430;

Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

YY 42 VOKKFTKWN 52
|||
79 VOKKFTKWN 89

DB 42 VOKKFTKWN 52
|||
79 VOKKFTKWN 89

RESULT 50

ID AAU14697 standard; protein; 5447 AA.

AC AAU14697;

DT 24-OCT-2001 (first entry)

DE Novel bone marrow polypeptide #96.

XX Bone marrow; diagnostic; therapeutic; gene therapy; antigenic;
XX haematopoiesis; myeloid; lymph cell disorder; tissue regeneration;
XX wound healing; nutritional supplement; immune disorder;

KM severe combined immunodeficiency; SCID.

XX Homo sapiens.

PN WO200157187-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US003782.

PR 03-FEB-2000; 2000US-00496914.

PR 20-JUN-2000; 2000US-00598075.

PR 19-JUL-2000; 2000US-00620325.

PR 30-NOV-2000; 2000US-0250683P.

XX (HYSE-) HYSEQ INC.

PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Zhou P, Xue AJ;

PI Ren F, Drmanac RT;

DR WPI; 2001-488875/53.

PT Nucleic acids encoding bone marrow polypeptides, useful in diagnostic and
PT gene therapy.

PS Claim 10; Page 124-127; 392pp; English.

XX AAU14602-AAU14794 represent novel bone marrow polypeptides of the
XX invention. The proteins and corresponding coding sequences may be used in
XX the prevention, diagnosis and treatment of diseases associated with
XX inappropriate bone marrow polypeptide expression. For example, to treat
XX disorders associated with decreased expression by rectifying mutations or
XX deletions in a patient's genome that affect the activity of the
XX polypeptides by expressing inactive proteins or to supplement the
XX patient's own production of the polypeptides. Additionally, the nucleic
XX acids may be used to produce the polypeptides, by inserting the nucleic
XX acids into a host cell and culturing the cell to express the protein. The
XX nucleic acid and its complementary sequences may also be used as DNA
XX probes in diagnostic assays to detect and quantitate the presence of
XX similar nucleic acid sequences in samples, and therefore which patients
XX may be in need of restorative therapy. The proteins may also be used as
XX antigens in the production of antibodies against bone marrow proteins and
XX in assays to identify modulators of their expression and activity. The
XX anti-bone marrow protein antibodies and antagonists may also be used to
XX down regulate expression and activity. The antibodies may also be used as

CC diagnostic agents for detecting the presence of the protein in samples
 CC (e.g. by enzyme linked immunosorbant assay (ELISA)). The proteins may be
 CC used to regulate haematopoiesis activity, and consequently in the
 CC treatment of myeloid or lymph cell disorders; in tissue regeneration,
 CC such as wound healing; as a nutritional supplement; and in treatment of
 CC immune disorders such as severe combined immunodeficiency (SCID)
 XX

SQ Sequence 5447 AA;

Query Match 5.0%; Score 11; DB 4; Length 5447;

Best Local Similarity 100.0%; Pred. NO. 0.67; Mismatches 0; Indels 0; Gaps 0;

OY 42 VOKKTFTKWN 52
 |||||
 Db 96 VOKKTFTKWN 106

Search completed: March 4, 2006, 07:06:15
 Job time : 242 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2006, 07:02:16 ; Search time 230 Seconds
(without alignments)
674.853 Million cell updates/sec

Title: US-10-695-994A-7
Perfect score: 220
Sequence: 1 MELORTSSISGPLSPATYTGQ.....QNYKNFNSRTASHDSHSGM 220

Scoring table: @MGO
Gapop 60.0 , Gapext 60.0

Searched: 2166443 seqs, 705528306 residues

Word size: 6

Total number of hits satisfying chosen parameters: 8624

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : UniProt.05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	220	100.0	220	Q922V0_MOUSE	Q922V0 mus musculus
2	133	60.5	155	Q8C509_MOUSE	Q8C509 mus musculus
3	133	60.5	2154	Q5S0L9_MOUSE	Q5S0L9 mus musculus
4	133	60.5	2155	Q81X99_HUMAN	Q81X99 homo sapien
5	124	56.4	2154	Q9QWJ7_MOUSE	Q9QWJ7 mus musculus
6	97	44.1	149	Q6V0G7_CANFA	Q6V0G7 canis faml
7	97	44.1	1290	Q5DTR4_MOUSE	Q5DTR4 mus musculus
8	97	44.1	2358	Q6XD99_RAT	Q6XD99 rattus norv
9	97	44.1	2363	SPTB2_MOUSE	SPTB2 mus musculus
10	97	44.1	2363	Q5S0L8_MOUSE	Q5S0L8 mus musculus
11	97	44.1	2364	SPTB2_HUMAN	SPTB2 homo sapien
12	97	44.1	2377	Q59ER3_HUMAN	Q59ER3 homo sapien
13	96	43.6	2314	Q53R99_HUMAN	Q53R99 homo sapien
14	66	30.0	977	Q502M3_BRARE	Q502M3 brachydanto
15	65	29.5	2235	Q4REB5_TETNG	Q4REB5 tetractodon n
16	62	28.2	332	Q9QWJ7_MOUSE	Q9QWJ7 mus sp. bet
17	62	28.2	497	Q8BOJ3_MOUSE	Q8BOJ3 mus musculus
18	60	27.3	48	Q8V165_RAT	Q8V165 rattus norv
19	48	21.8	48	Q9RLY9_MOUSE	Q9RLY9 mus musculus
20	47	21.4	2275	Q4SR66_TETNG	Q4SR66 tetractodon n
21	46	20.9	2307	Q4SUVE_TETNG	Q4SUVE tetractodon n
22	46	20.9	2357	Q9DGM6_BRARE	Q9DGM6 brachydanto
23	38	17.3	2127	SPTB1_MOUSE	SPTB1 mus musculus
24	38	17.3	2136	SPTB1_HUMAN	SPTB1 homo sapien
25	38	17.3	2137	Q6XDA0_RAT	Q6XDA0 rattus norv
26	38	17.3	2332	Q59RPS_HUMAN	Q59RPS homo sapien
27	37	16.8	67	Q8WYB3_HUMAN	Q8WYB3 homo sapien
28	37	16.8	82	Q9WUX0_RAT	Q9WUX0 rattus norv
29	32	14.5	516	Q5D002_RAT	Q5D002 rattus norv
30	28	12.7	2413	Q4SF10_TETNG	Q4SF10 tetractodon n
31	27	12.3	794	Q5XG36_XENILA	Q5XG36 xenopus lae

32	27	12.3	1060	Q7ZX11_XENILA	Q7ZX11 xenopus lae
33	27	12.3	2388	SPTN2_RAT	SPTN2 rattus norv
34	27	12.3	2388	Q68FG2_MOUSE	Q68FG2 mus musculus
35	27	12.3	2390	SPTN2_HUMAN	SPTN2 homo sapien
36	25	11.4	51	Q71VG2_HUMAN	Q71VG2 homo sapien
37	22	10.0	2002	Q71S06_HUMAN	Q71S06 homo sapien
38	22	10.0	2555	Q91Z66_MOUSE	Q91Z66 mus musculus
39	22	10.0	2561	Q8V165_MOUSE	Q8V165 mus musculus
40	22	10.0	2564	SPTN4_HUMAN	SPTN4 homo sapien
41	22	10.0	2564	Q71S07_HUMAN	Q71S07 homo sapien
42	19	8.6	1410	Q6NNX2_DROME	Q6NNX2 drosophila
43	19	8.6	2291	SPTCB_DROME	SPTCB drosophila
44	17	7.7	20	Q9TR82_PIG	Q9TR82 sus scrofa
45	16	7.3	707	Q28297_CANFA	Q28297 canis faml
46	16	7.3	2296	Q4SPH4_ANOGA	Q4SPH4 anophles g
47	15	6.8	129	Q4RFW3_TETNG	Q4RFW3 tetractodon n
48	14	6.4	2257	Q9U9J7_CABEL	Q9U9J7 caenorhabdi
49	14	6.4	2299	Q6WPE8_CABEL	Q6WPE8 caenorhabdi
50	14	6.4	2302	Q95ZL8_CABEL	Q95ZL8 caenorhabdi
51	14	6.4	2302	Q9U9J8_CABEL	Q9U9J8 caenorhabdi
52	13	6.4	2339	Q961X7_BRUMA	Q961X7 briglia mala
53	13	5.9	373	Q4SR67_TETNG	Q4SR67 tetractodon n
54	12	5.5	1315	Q4S2Y7_TETNG	Q4S2Y7 tetractodon n
55	12	5.5	3641	Q7ZTH4_FUGRU	Q7ZTH4 fugu rubrip
56	11	5.0	20	Q8WYF6_HUMAN	Q8WYF6 homo sapien
57	11	5.0	41	Q5JXB4_HUMAN	Q5JXB4 homo sapien
58	11	5.0	50	Q8WYX4_HUMAN	Q8WYX4 homo sapien
59	11	5.0	65	Q810C9_RAT	Q810C9 rattus norv
60	11	5.0	101	Q14172_HUMAN	Q14172 homo sapien
61	11	5.0	108	Q8C221_MOUSE	Q8C221 mus musculus
62	11	5.0	256	Q816B6_DROME	Q816B6 drosophila
63	11	5.0	385	Q13696_HUMAN	Q13696 homo sapien
64	11	5.0	420	Q516B1_PHTHI	Q516B1 entameoba h
65	11	5.0	440	Q5R9D0_PONPY	Q5R9D0 pongo pygma
66	11	5.0	462	Q6PK12_HUMAN	Q6PK12 homo sapien
67	11	5.0	470	Q5R8S0_PONPY	Q5R8S0 pongo pygma
68	11	5.0	565	Q8ATW7_ENCCPU	Q8ATW7 encephalito
69	11	5.0	676	Q8C8S5_MOUSE	Q8C8S5 mus musculus
70	11	5.0	719	Q5JXW4_HUMAN	Q5JXW4 homo sapien
71	11	5.0	772	Q5JXQ4_HUMAN	Q5JXQ4 homo sapien
72	11	5.0	786	Q503P9_BRARE	Q503P9 brachydanto
73	11	5.0	964	PLEC1_MOUSE	PLEC1 mus musculus
74	11	5.0	1375	Q4RUK6_TETNG	Q4RUK6 tetractodon n
75	11	5.0	1445	Q4SY67_TETNG	Q4SY67 tetractodon n
76	11	5.0	1454	Q7ZMK4_XENILA	Q7ZMK4 xenopus lae
77	11	5.0	1697	Q4SY99_TETNG	Q4SY99 tetractodon n
78	11	5.0	3322	Q6BEO3_CABEL	Q6BEO3 caenorhabdi
79	11	5.0	3323	Q6BEO4_CABEL	Q6BEO4 caenorhabdi
80	11	5.0	3405	Q6BEO5_CABEL	Q6BEO5 caenorhabdi
81	11	5.0	3436	Q6BEP8_CABEL	Q6BEP8 caenorhabdi
82	11	5.0	3436	Q6BEP9_CABEL	Q6BEP9 caenorhabdi
83	11	5.0	3436	Q18290_CABEL	Q18290 caenorhabdi
84	11	5.0	3522	Q86NF7_CABEL	Q86NF7 caenorhabdi
85	11	5.0	3680	DMD_CANFA	DMD canis faml
86	11	5.0	3685	Q7KZ48_HUMAN	Q7KZ48 homo sapien
87	11	5.0	3685	Q5UYU0_HUMAN	Q5UYU0 homo sapien
88	11	5.0	3685	Q6EUT7_CABEL	Q6EUT7 caenorhabdi
89	11	5.0	4410	Q6S388_MOUSE	Q6S388 mus musculus
90	11	5.0	4511	Q6S399_RAT	Q6S399 rattus norv
91	11	5.0	4513	Q6S379_HUMAN	Q6S379 homo sapien
92	11	5.0	4521	Q6S392_MOUSE	Q6S392 mus musculus
93	11	5.0	4523	Q6S3A3_RAT	Q6S3A3 rattus norv
94	11	5.0	4525	Q6S381_HUMAN	Q6S381 homo sapien
95	11	5.0	4533	Q6S382_HUMAN	Q6S382 homo sapien
96	11	5.0	4534	Q6S393_MOUSE	Q6S393 mus musculus
97	11	5.0	4536	Q6S3A4_RAT	Q6S3A4 rattus norv
98	11	5.0	4543	Q6S384_MOUSE	Q6S384 mus musculus
99	11	5.0	4543	Q6S387_MOUSE	Q6S387 mus musculus
100	11	5.0	4545	Q6S395_RAT	Q6S395 rattus norv
101	11	5.0	4545	Q6S398_RAT	Q6S398 rattus norv
102	11	5.0	4547	Q6S376_HUMAN	Q6S376 homo sapien
103	11	5.0	4547	Q6S378_HUMAN	Q6S378 homo sapien
104	11	5.0	4547	Q6S378_HUMAN	Q6S378 homo sapien

105	11	5.0	4550	2	066385 MOUSE	066385 mus musculus	178	8	3.6	440	2	050NF1_ENTHI	050NF1 entamoeba h
106	11	5.0	4551	2	066377_HUMAN	066377 homo sapien	179	8	3.6	457	2	050X88_ENTHI	050X88 entamoeba h
107	11	5.0	4552	2	066396_RAT	066396 rattus norv	180	8	3.6	472	2	04R0U7_TETNG	04R0U7 tetraodon n
108	11	5.0	4572	2	066394_MOUSE	066394 mus musculu	181	8	3.6	484	2	050W17_ENTHI	050W17 entamoeba h
109	11	5.0	4574	2	066383_HUMAN	066383 homo sapien	182	8	3.6	512	2	050Z94_ENTHI	050Z94 entamoeba h
110	11	5.0	4574	2	066345_RAT	066345 rattus norv	183	8	3.6	534	2	0518F3_ENTHI	0518F3 entamoeba h
111	11	5.0	4684	2	066380_HUMAN	066380 homo sapien	184	8	3.6	534	2	04RNG2_TETNG	04RNG2 tetraodon n
112	11	5.0	4686	2	066389_MOUSE	066389 mus musculu	185	8	3.6	572	2	04W116_ASPFU	04W116 aspeygiilus
113	11	5.0	4688	2	066340_RAT	066340 rattus norv	186	8	3.6	592	2	004096_ARATH	004096 arabidopsis
114	11	5.0	4944	2	07K7B8_CABEL	07K7B8 caenorhabdi	187	8	3.6	605	2	07T3H6_BRARE	07T3H6 brachydanio
115	11	5.0	4955	2	0867D9_CABEL	0867D9 caenorhabdi	188	8	3.6	619	2	0514P9_ENTHI	0514P9 entamoeba h
116	11	5.0	5327	1	MACF1_MOUSE	094XZ0 mus musculu	189	8	3.6	737	2	050SC7_ENTHI	050SC7 entamoeba h
117	11	5.0	5430	2	MACF1_HUMAN	094XZ0 mus musculu	190	8	3.6	737	2	0514A0_MOUSE	0514A0 mus musculu
118	11	5.0	5430	2	04PL15_MOUSE	04PL15 mus musculu	191	8	3.6	762	1	PD0C8_DROME	PD0C8 drome
119	11	5.0	7735	2	09VMT8_DROME	09VMT8 drosophila	192	8	3.6	739	2	050RA3_ENTHI	050RA3 entamoeba h
120	11	5.0	7735	2	09VMT8_DROME	09VMT8 drosophila	193	8	3.6	762	2	080016_FUGRU	080016 fugu rubrip
121	10	4.5	1064	2	0965L0_CABEL	0965L0 caenorhabdi	194	8	3.6	822	2	09N5P5_CAREL	09N5P5 caenorhabdi
122	10	4.5	1844	2	05BDW2_EMBENI	05BDW2 aspeygiilus	195	8	3.6	839	2	06DES1_XENLA	06DES1 xenopus lae
123	10	4.5	2285	2	07KAJ2_DROME	07KAJ2 drosophila	196	8	3.6	857	2	050UD2_ENTHI	050UD2 entamoeba h
124	10	4.5	2396	2	077291_DROME	077291 drosophila	197	8	3.6	857	2	051GH3_ENTHI	051GH3 entamoeba h
125	10	4.5	5201	2	07KJN8_DROME	07KJN8 drosophila	198	8	3.6	858	2	P90545_ENTHI	P90545 entamoeba h
126	10	4.5	5201	2	08ML80_DROME	08ML80 drosophila	199	8	3.6	887	2	04PA36_USITMA	04PA36 uscitlago ma
127	10	4.5	5385	2	09V6V3_DROME	09V6V3 drosophila	200	8	3.6	889	2	04RZ26_TETNG	04RZ26 tetraodon n
128	10	4.5	5390	2	09V6V4_DROME	09V6V4 drosophila	201	8	3.6	890	2	06NRW6_XENLA	06NRW6 xenopus lae
129	10	4.5	5501	2	09V6V2_DROME	09V6V2 drosophila	202	8	3.6	890	2	06P0U5_BRARE	06P0U5 brachydanio
130	10	4.1	5501	2	0811A1_9MURI	0811A1 mus sp. dys	203	8	3.6	892	2	062744_RAT	062744 rattus norv
131	9	4.1	47	2	05THP1_HUMAN	05THP1 homo sapien	204	8	3.6	894	1	ACTN2_HUMAN	ACTN2 human
132	9	4.1	99	2	063770_9MURI	063770 rattus sp.	205	8	3.6	894	2	ACTN2_MOUSE	ACTN2 mouse
133	9	4.1	107	2	063771_9MURI	063771 rattus sp.	206	8	3.6	894	2	05VXF1_HUMAN	05VXF1 human
134	9	4.1	124	2	04RQ46_TETNG	04RQ46 tetraodon n	207	8	3.6	894	2	05FWT5_MOUSE	05FWT5 mouse
135	9	4.1	153	2	05SYX1_HUMAN	05SYX1 homo sapien	208	8	3.6	894	2	06R3Q4_MOUSE	06R3Q4 mouse
136	9	4.1	269	2	08C7R5_MOUSE	08C7R5 mus musculu	209	8	3.6	894	2	06DF05_XENTR	06DF05 xenopus tro
137	9	4.1	292	1	TNNI1_CHTNI	07MY33 chlamey nlp	210	8	3.6	896	2	06IP14_XENLA	06IP14 xenopus lae
138	9	4.1	293	2	015988_PATYE	015988 patinopecte	211	8	3.6	896	2	06DFU3_XENLA	06DFU3 xenopus lae
139	9	4.1	314	2	015987_PATYE	015987 patinopecte	212	8	3.6	896	2	07SYD3_BRARE	07SYD3 brachydanio
140	9	4.1	589	2	06YBR7_BRARE	06YBR7 brachydanio	213	8	3.6	896	2	06TNW2_BRARE	06TNW2 brachydanio
141	9	4.1	589	2	05RCK8_PONPY	05RCK8 pongo pygma	214	8	3.6	897	1	ACTN2_CHICK	ACTN2 chick
142	9	4.1	603	2	09VOM7_PVRAB	09VOM7 pyrococcus	215	8	3.6	897	2	04ST30_TETNG	04ST30 tetraodon n
143	9	4.1	603	2	04IFB0_GIBZE	04IFB0 gibberella	216	8	3.6	898	2	08AX99_BRARE	08AX99 brachydanio
144	9	4.1	633	2	04IFB0_GIBZE	04IFB0 gibberella	217	8	3.6	901	2	07SYE2_BRARE	07SYE2 brachydanio
145	9	4.1	689	2	05T0V7_HUMAN	05T0V7 homo sapien	218	8	3.6	904	1	ACTN4_CHICK	ACTN4 chick
146	9	4.1	732	2	0810I5_MOUSE	0810I5 mus musculu	219	8	3.6	904	2	06DCS8_XENLA	06DCS8 xenopus lae
147	9	4.1	732	2	06PEL2_MOUSE	06PEL2 mus musculu	220	8	3.6	904	2	06DPP7_XENTR	06DPP7 xenopus tro
148	9	4.1	874	2	018109_CABEL	018109 caenorhabdi	221	8	3.6	911	1	ACTN4_HUMAN	ACTN4 human
149	9	4.1	1086	2	05ZME3_CHICK	05ZME3 gallus gall	222	8	3.6	911	2	05RCS6_PONPY	05RCS6 pongo pygma
150	9	4.1	1304	2	09PWF6_MORSA	09PWF6 morone saxa	223	8	3.6	911	2	06P786_RAT	06P786 rattus norv
151	9	4.1	1678	1	BPABR_MOUSE	060824 mus musculu	224	8	3.6	912	2	ACTN4_MOUSE	ACTN4 mouse
152	9	4.1	2063	2	04S6Z2_TETNG	04S6Z2 tetraodon n	225	8	3.6	918	2	05LJPI_DROME	05LJPI drome
153	9	4.1	3029	2	04S6Z2_TETNG	04S6Z2 tetraodon n	226	8	3.6	965	2	05H724_FUGRU	05H724 fugu rubrip
154	9	4.1	3214	1	BPAL_HUMAN	055147 rattus norv	227	8	3.6	1011	1	PERO1_MOUSE	PERO1 mouse
155	9	4.1	3419	2	055147_RAT	055147 rattus norv	228	8	3.6	1035	2	0677W7_HUMAN	0677W7 human
156	9	4.1	3429	2	008614_MOUSE	008614 mus musculu	229	8	3.6	1043	2	0617W8_MOUSE	0617W8 mouse
157	9	4.1	3432	2	05JCF5_CANPA	05JCF5 canis famli	230	8	3.6	1089	2	05X908_STRCO	05X908 streptomyce
158	9	4.1	3433	1	UTRO_HUMAN	P46939 homo sapien	231	8	3.6	1172	2	07R3Q6_GIALA	07R3Q6 giardia lam
159	9	4.1	3527	2	05S2S7_HUMAN	05S2S7 homo sapien	232	8	3.6	1190	2	09HS95_HALSA	09HS95 halobacteri
160	9	4.1	3535	2	07T1P2_FUGRU	07T1P2 fugu rubrip	233	8	3.6	1253	2	0729R3_DESVH	0729R3 desulfovibr
161	9	4.1	3650	1	DMD_CHICK	P11533 gallus gall	234	8	3.6	1462	2	04SXZ7_TETNG	04SXZ7 tetraodon n
162	9	4.1	3674	1	DMD_MOUSE	05GN48 sus scrofa	235	8	3.6	1553	2	07Z6C1_HUMAN	07Z6C1 homo sapien
163	9	4.1	3678	1	DMD_MOUSE	P11531 mus musculu	236	8	3.6	1609	2	07RYS2_NEUCR	07RYS2 neurospora
164	9	4.1	5457	1	BPAL_MOUSE	05TBE1 homo sapien	237	8	3.6	1645	2	099140_DROME	099140 drosophila
165	9	4.1	7389	1	BPAL_MOUSE	05TBE1 mus musculu	238	8	3.6	1648	2	06FES4_CANGA	06FES4 candida gla
166	9	4.1	7543	2	05TBT2_HUMAN	05TBT2 homo sapien	239	8	3.6	1738	2	051ED7_ENTHI	051ED7 entamoeba h
167	8	3.6	72	2	061063_MOUSE	061063 mus musculu	240	8	3.6	1843	2	04ARR8_TETNG	04ARR8 tetraodon n
168	8	3.6	107	2	08RLJ1_PROKE	08RLJ1 providencia	241	8	3.6	1843	2	096819_CAREL	096819 caenorhabdi
169	8	3.6	110	2	08A9W5_SALSA	08A9W5 salmo salar	242	8	3.6	1846	2	016131_CABEL	016131 caenorhabdi
170	8	3.6	169	2	08F493_LEPIC	08F493 leprospira	243	8	3.6	2053	2	0694W8_XENIA	0694W8 xenopus lae
171	8	3.6	169	2	08F493_LEPIC	08F493 leprospira	244	8	3.6	2221	2	07S3C3_NEUCR	07S3C3 neurospora
172	8	3.6	246	2	07XP04_ORYSA	07XP04 oryza sativ	245	8	3.6	2414	1	EP3100_HUMAN	EP3100 human
173	8	3.6	301	2	05QMA0_ORYSA	05QMA0 oryza sativ	246	8	3.6	2414	2	05W1B6_HUMAN	05W1B6 homo sapien
174	8	3.6	417	2	09PUC1_BRARE	09PUC1 brachydanio	247	8	3.6	2539	2	04S4B6_TETNG	04S4B6 tetraodon n
175	8	3.6	418	2	06NV30_BRARE	06NV30 brachydanio	248	8	3.6	2730	2	04SVN3_TETNG	04SVN3 tetraodon n
176	8	3.6	418	2	06PEZ6_BRARE	06PEZ6 brachydanio	249	8	3.6	3497	1	DMDA_DROME	DMDA drosophila
177	8	3.6	419	2	050TP1_ENTHI	050TP1 entamoeba h	250	8	3.6				

251	8	3.6	4052	2	Q70A48	ANOXA	Q79a48	anopheles	g	324	7	3.2	244	2	Q9RAE5	RHILV	Q9rae5	rhizobium	1
252	8	3.6	4061	2	Q60M30	CABER	Q60m30	caenorhabdi		325	7	3.2	247	2	Q7Z617	HUMAN	Q7ze617	homo sapien	
253	8	3.6	4063	2	Q02A25	CABEL	Q02a25	caenorhabdi		326	7	3.2	251	2	Q7Z617	SALTY	Q7ze617	salmonella	
254	8	3.6	4097	2	Q9VZ03	DROME	Q9vz03	drosophila		327	7	3.2	251	2	Q7CPB8	SALTY	Q7cpb8	salmonella	
255	8	3.6	4118	2	Q7KV65	DROME	Q7kv65	drosophila		328	7	3.2	251	2	Q8XP90	SALTY	Q8xp90	salmonella	
256	8	3.6	4207	2	Q7KV70	DROME	Q7kv70	drosophila		329	7	3.2	251	2	Q5PKV4	SALTY	Q5pkv4	salmonella	
257	7	3.2	31	2	Q8WYF3	HUMAN	Q8wyf3	homo sapien		330	7	3.2	255	2	Q8WYF6	HUMAN	Q8wyf6	homo sapien	
258	7	3.2	41	2	Q9TYX7	9BIVA	Q9tyx7	mytilus	ca	331	7	3.2	266	1	PENKI	BACSU	Q9tyx7	mytilus	ca
259	7	3.2	42	2	Q9TY10	PIAFA	Q9ty10	plasmidum		332	7	3.2	267	2	Q70Q37	GRALA	Q70q37	gratidia	1m
260	7	3.2	44	2	Q9TSS4	BOVIN	Q9tss4	bos taurus		333	7	3.2	267	2	Q57652	PUGRI	Q57652	lugu rubrip	
261	7	3.2	46	2	Q3E214	9HIVI	Q3e214	human	immun	334	7	3.2	268	2	Q6K2H1	ORYSA	Q6k2h1	oryza sativ	
262	7	3.2	56	2	Q7X9L5	WHEAT	Q7x9l5	tritilicum	ae	335	7	3.2	269	1	MIND	GUTH	Q5z316	nocarpa	fa
263	7	3.2	76	2	Q5BYC6	SCHJA	Q5byc6	schistosoma		336	7	3.2	270	2	Q5Z316	NOCFA	Q5z316	nocarpa	fa
264	7	3.2	87	2	Q9RUF7	LOTICO	Q9ruf7	lotus corni		337	7	3.2	272	2	Q9LEB9	JUGNT	Q9leb9	juglans nig	
265	7	3.2	93	2	Q9SDP9	PHYCP	Q9sdp9	phytophthor		338	7	3.2	273	2	Q4V1B7	BACCC	Q4v1b7	bacillus ce	
266	7	3.2	95	2	Q5UWU0	HALMA	Q5uwu0	halotartula		339	7	3.2	278	2	Q7PUB4	ANOXA	Q7pue4	anopheles	g
267	7	3.2	102	2	Q50TW8	ENTHI	Q50tw8	entamoeba	h	340	7	3.2	279	2	Q6AHO3	CABEL	Q6ahq3	caenorhabdi	
268	7	3.2	107	2	Q5P1N6	AZOSE	Q5p1n6	azocarcus	sp	341	7	3.2	279	2	Q4Q281	LEIMA	Q4q281	leishmania	
269	7	3.2	115	2	Q829B7	STRAM	Q829b7	streptomyce		342	7	3.2	279	2	Q97FC2	CIOAB	Q97fc2	clostridium	
270	7	3.2	116	2	Q5FPT8	GLUOX	Q5fpt8	glucosinolat		343	7	3.2	282	2	Q6ACE8	9ARCH	Q6ace8	uncultured	
271	7	3.2	123	2	Q4I284	GIBZE	Q4i284	gibberella		344	7	3.2	282	2	Q5YOM1	9ARCH	Q5yom1	uncultured	
272	7	3.2	123	2	Q4IUT8	AZOV	Q4iut8	azocobacter		345	7	3.2	284	2	Q57VU1	9TRYP	Q57vj1	trypanosoma	
273	7	3.2	125	2	Q14I74	HUMAN	Q14i74	homo sapien		346	7	3.2	290	2	Q9FWX9	ARATH	Q9fwx9	arabidopsis	
274	7	3.2	133	2	Q4ZCP3	9CAUD	Q4zcp3	bacteriophag		347	7	3.2	291	2	Q9M140	ARATH	Q9m140	arabidopsis	
275	7	3.2	133	2	Q6GAN6	STPAS	Q6gan6	staphylococ		348	7	3.2	292	2	Q8YTY9	ANASP	Q8yty9	anabaena	sp
276	7	3.2	133	2	Q6GAN1	STPAR	Q6gan1	staphylococ		349	7	3.2	296	2	Q8YHN9	MOUSE	Q8yhn9	mus musculu	
277	7	3.2	133	2	Q8WVH5	STRAR	Q8wvh5	staphylococ		350	7	3.2	297	2	Q7Z4C7	HUMAN	Q7z4c7	homo sapien	
278	7	3.2	133	2	Q6DHG4	BRARE	Q6dhg4	brachydanio		351	7	3.2	300	2	Q8TLE1	METAC	Q8tle1	methanosarc	
279	7	3.2	134	2	Q640X9	XENIA	Q640x9	xenopus	lae	352	7	3.2	301	2	Q4SF76	TEING	Q4sf76	tetrahodon	n
280	7	3.2	141	2	Q6ELF9	CUCOSA	Q6elf9	cucumis	bat	353	7	3.2	309	2	Q6ZVW9	HUMAN	Q6zvw9	homo sapien	
281	7	3.2	150	2	Q70FEM4	ARALP	Q70fem4	arabidopsis		354	7	3.2	310	2	Q50MD9	ENTHI	Q50md9	entamoeba	h
282	7	3.2	150	2	Q70FEM4	ARALP	Q70fem4	arabidopsis		355	7	3.2	311	2	Q88016	STRCO	Q88016	streptococ	
283	7	3.2	150	2	Q70FEM1	ARALP	Q70fem1	arabidopsis		356	7	3.2	313	2	Q8W596	WEDSA	Q8w596	medicago	sa
284	7	3.2	150	2	Q70FNO	ARALP	Q70fno	arabidopsis		357	7	3.2	314	2	Q94436	9BIVA	Q94436	chlamys	nip
285	7	3.2	152	2	Q755M6	ASHGA	Q755m6	asbhyia	goss	358	7	3.2	314	2	Q5ZH19	PSEME	Q5zh19	pseudomonas	
286	7	3.2	154	2	Q87PJ4	VIBPA	Q87pj4	vibriol	para	359	7	3.2	316	2	Q85470	STRPY	Q85470	streptococ	
287	7	3.2	155	2	Q5FE209	MOUSE	Q5fe209	mus musculu		360	7	3.2	316	2	Q5XCZ9	STRPY	Q5xcz9	streptococ	
288	7	3.2	158	2	Q8CFI8	MOUSE	Q8cfi8	mus musculu		361	7	3.2	316	2	Q8K838	STRPY	Q8k838	streptococ	
289	7	3.2	162	2	Q4V739	GYIRU	Q4v739	rice	tungro	362	7	3.2	319	2	Q15985	PATVE	Q15985	patinopecte	
290	7	3.2	163	2	Q7XQ51	ORYSA	Q7xq51	oryza sativ		363	7	3.2	321	2	Q15986	PATVE	Q15986	patinopecte	
291	7	3.2	172	1	LSM1	YEAST	P47017	saccharomyc		364	7	3.2	322	2	Q4FUC5	9GAMA	Q4fuc5	psychobact	
292	7	3.2	174	2	Q4UGV8	THEAN	Q4ugv8	theileria	a	365	7	3.2	324	2	Q8PP4	ORYSA	Q8pp4	oryza sativ	
293	7	3.2	174	2	Q4UGV8	THEAN	Q4ugv8	theileria	a	366	7	3.2	325	2	Q09144	PATVE	Q09144	patinopecte	
294	7	3.2	175	2	Q6GCM7	GYIRU	Q6gcm7	infectious		367	7	3.2	327	2	Q75FZ9	LEPIC	Q75fz9	leptospira	
295	7	3.2	175	2	Q6GCM7	GYIRU	Q6gcm7	infectious		368	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
296	7	3.2	175	2	Q6GCM7	GYIRU	Q6gcm7	infectious		369	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
297	7	3.2	175	2	Q6GCM7	GYIRU	Q6gcm7	infectious		370	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
298	7	3.2	175	2	Q6GCM7	GYIRU	Q6gcm7	infectious		371	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
299	7	3.2	175	2	Q6GCM7	GYIRU	Q6gcm7	infectious		372	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
300	7	3.2	175	2	Q6GCM7	GYIRU	Q6gcm7	infectious		373	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
301	7	3.2	175	2	Q6GCM7	GYIRU	Q6gcm7	infectious		374	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
302	7	3.2	179	2	Q6GCM7	GYIRU	Q6gcm7	infectious		375	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
303	7	3.2	182	2	Q5QXK6	IDILO	Q5qxk6	idionaria		376	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
304	7	3.2	184	2	Q9TUP8	PIC	Q9tup8	scrofa		377	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
305	7	3.2	185	2	Q7NM59	GLOVI	Q7nm59	gloeobacter		378	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
306	7	3.2	186	2	Q8W220	HUMAN	Q8w220	homo sapien		379	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
307	7	3.2	186	2	Q8W220	HUMAN	Q8w220	homo sapien		380	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
308	7	3.2	191	2	Q84S95	RAPSA	Q84s95	rapphanus	sa	381	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
309	7	3.2	211	2	Q4P4V2	USTMA	Q4p4v2	ustilago	ma	382	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
310	7	3.2	213	2	Q8EMG7	OCEIH	Q8emg7	oceanobact		383	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
311	7	3.2	215	2	Q8TCN6	HUMAN	Q8tcn6	homo sapien		384	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
312	7	3.2	215	2	Q8TCN6	HUMAN	Q8tcn6	homo sapien		385	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
313	7	3.2	217	2	Q7Z2X0	HUMAN	Q7z2x0	homo sapien		386	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
314	7	3.2	218	2	Q9H5W4	HALSA	Q9h5w4	halobacteri		387	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
315	7	3.2	224	2	Q7NM08	CHRYO	Q7nm08	chromobact		388	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
316	7	3.2	231	2	Q5BWK9	SACOF	Q5bwk9	saccharom	o	389	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
317	7	3.2	231	2	Q5BWK9	SACOF	Q5bwk9	saccharom	o	390	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
318	7	3.2	233	2	Q83G85	STRAM	Q83g85	streptomyce		391	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
319	7	3.2	233	2	Q7Z9K9	DSVH	Q7z9k9	desulfovibr		392	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
320	7	3.2	237	1	GIDB	MYCPE	Q8euw9	mycoplasma		393	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
321	7	3.2	237	1	GIDB	MYCPE	Q8euw9	mycoplasma		394	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
322	7	3.2	237	1	GIDB	MYCPE	Q8euw9	mycoplasma		395	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	
323	7	3.2	244	2	P72285	RHILE	P72285	rhizobium	1	396	7	3.2	327	2	Q8EY39	LEPIN	Q8ey39	leptospira	

397	7	3.2	356	1	FL3H_CALCH	Q05963	callistephu	470	7	3.2	419	2	Q6UD79_ONCMY	Q6UD79	oncothynehu
398	7	3.2	357	1	FL3H_MATIN	Q05965	mathioloia i	471	7	3.2	422	2	Q4WVZ8_ASPFU	Q4WVZ8	aspeygiilus
399	7	3.2	357	2	Q6VZ3_HUMAN	Q6VZ3	homo sapien	472	7	3.2	423	2	Q8ROA3_MOUSE	Q8ROA3	mouse
400	7	3.2	357	2	Q5SQS_HUMAN	Q5SQS	homo sapien	473	7	3.2	423	2	Q7SZM3_ICTPU	Q7SZM3	ictalurus p
401	7	3.2	357	2	Q49982_9ASTR	Q49982	chrysanthem	474	7	3.2	428	2	Q6UJ36_CABER	Q6UJ36	caenothabdi
402	7	3.2	357	2	Q40344_MBSA	Q40344	medicago sa	475	7	3.2	428	2	Q808D4_VYIRU	Q808D4	caenothabdi
403	7	3.2	357	2	Q42927_MEDSA	Q42927	medicago sa	476	7	3.2	428	2	Q808D5_VYIRU	Q808D5	caenothabdi
404	7	3.2	357	2	Q4QUR9_DAUCA	Q4QUR9	daucus caro	477	7	3.2	431	2	Q5SOS7_HUMAN	Q5SOS7	homo sapien
405	7	3.2	357	2	Q9X221_THEMA	Q9X221	thermoca	478	7	3.2	433	2	Q22093_CABEL	Q22093	caenothabdi
406	7	3.2	358	1	FL3H_ARATH	Q9A818	arabidopsis	479	7	3.2	442	2	Q808U0_VYIRU	Q808U0	infectious
407	7	3.2	358	1	Q56180_GSOLN	Q56180	solanium pin	480	7	3.2	450	2	Q65106_BACLD	Q65106	baclillus l1
408	7	3.2	358	2	Q5DWM0_ARATH	Q5DWM0	arabidopsis	481	7	3.2	451	2	Q4T9X0_TETNG	Q4T9X0	tetradon n
409	7	3.2	358	2	Q818C8_SOLTU	Q818C8	solanium tub	482	7	3.2	453	2	Q5M17_CRYNE	Q5M17	cryne
410	7	3.2	359	2	Q6SG19_9BACT	Q6SG19	uncultured	483	7	3.2	453	2	Q5KJ77_CRYNE	Q5KJ77	cryne
411	7	3.2	359	2	Q6UCT3_9GAMM	Q6UCT3	uncultured	484	7	3.2	455	2	Q9GTW7_DROYA	Q9GTW7	droya
412	7	3.2	361	2	Q7SZ61_BRARE	Q7SZ61	brachydantio	485	7	3.2	460	2	Q51F81_ENTHI	Q51F81	entameoba h
413	7	3.2	362	2	Q9ZMR0_CITST	Q9ZMR0	citrus sine	486	7	3.2	465	2	Q51U30_MAGGR	Q51U30	magnaporthe
414	7	3.2	363	2	Q6BEH2_EUSGR	Q6BEH2	eubetoma gra	487	7	3.2	468	2	Q9H6A0_HUMAN	Q9H6A0	homo sapien
415	7	3.2	364	1	FL3H_VITVI	Q6DBH2	vitis vinif	488	7	3.2	468	2	Q5T5V6_HUMAN	Q5T5V6	homo sapien
416	7	3.2	364	2	Q66ME9_FRAAN	Q66ME9	fragaria an	489	7	3.2	470	2	Q9T0C4_ARATH	Q9T0C4	arabidopsis
417	7	3.2	364	2	Q66MFO_FRAAN	Q66MFO	fragaria an	490	7	3.2	471	2	Q9BSU0_HUMAN	Q9BSU0	homo sapien
418	7	3.2	365	2	Q59169_GENTR	Q59169	gentiana tr	491	7	3.2	472	2	Q5CZV9_BRARE	Q5CZV9	brachydantio
419	7	3.2	365	2	Q59170_GENTR	Q59170	gentiana tr	492	7	3.2	476	2	Q4YT27_PLABE	Q4YT27	plasmodium
420	7	3.2	365	2	Q4QUS0_9APIA	Q4QUS0	pimpinella	493	7	3.2	478	2	Q97TX5_SUTSO	Q97TX5	sutso
421	7	3.2	365	2	Q4QUR7_CONMC	Q4QUR7	conium macu	494	7	3.2	481	2	Q54072_SACER	Q54072	sacer
422	7	3.2	365	2	Q7NOL3_CHRYO	Q7NOL3	chromobacte	495	7	3.2	483	2	Q62WQ0_MOUSE	Q62WQ0	mouse
423	7	3.2	365	2	Q23825_IPONI	Q23825	ipomoea nil	496	7	3.2	483	2	Q61WL3_BRARE	Q61WL3	brare
424	7	3.2	366	2	Q23825_IPONI	Q23825	ipomoea nil	497	7	3.2	488	2	Q5OLM0_ENTHI	Q5OLM0	entameoba h
425	7	3.2	366	2	Q9SP53_DAUCA	Q9SP53	daucus caro	498	7	3.2	496	2	Q4MLJ8_BACCE	Q4MLJ8	baclillus ce
426	7	3.2	366	2	Q4QUR5_9APIA	Q4QUR5	anethum gra	499	7	3.2	498	2	Q56776_XANCA	Q56776	xanthomonas
427	7	3.2	367	2	Q8LP21_9SOLA	Q8LP21	nietherbergi	500	7	3.2	498	2	Q9QSU8_9ITVI	Q9QSU8	9itvi
428	7	3.2	367	2	Q96472_IPOPU	Q96472	ipomoea pur	501	7	3.2	499	2	Q4UM41_XANCP	Q4UM41	xanthomonas
429	7	3.2	368	2	Q6DV45_CAMSI	Q6DV45	camellia si	502	7	3.2	499	2	Q8P805_XANCP	Q8P805	xanthomonas
430	7	3.2	368	2	Q7XZQ7_PETCR	Q7XZQ7	petroselinu	503	7	3.2	500	2	Q5A2J1_CANAL	Q5A2J1	candida alb
431	7	3.2	368	2	Q9SX85_IPOBA	Q9SX85	ipomoea bat	504	7	3.2	501	2	Q8DUU0_STRMU	Q8DUU0	strepococc
432	7	3.2	368	2	Q9ZQ57_IPOBA	Q9ZQ57	ipomoea bat	505	7	3.2	502	2	Q4RYA3_TETNG	Q4RYA3	tetradon n
433	7	3.2	368	2	Q9ZQ58_IPOBA	Q9ZQ58	ipomoea bat	506	7	3.2	505	2	Q96DL0_HUMAN	Q96DL0	homo sapien
434	7	3.2	368	2	Q4QUR6_9APIA	Q4QUR6	ammaj majus	507	7	3.2	507	2	Q810Q4_EMENT	Q810Q4	emeritella
435	7	3.2	370	2	Q4VBR4_HUMAN	Q4VBR4	homo sapien	508	7	3.2	507	2	Q5KKH8_CRYNE	Q5KKH8	cryptococcu
436	7	3.2	370	2	Q5ZH06_PSEPS	Q5ZH06	pseudomonas	509	7	3.2	507	2	Q55VK0_CRYNE	Q55VK0	cryptococcu
437	7	3.2	372	2	Q6EPA3_9LAMI	Q6EPA3	homo sapien	510	7	3.2	508	2	Q6BK97_DBRBA	Q6BK97	debrayomyce
438	7	3.2	376	1	MJD1_HUMAN	Q9A874	oryza sativ	511	7	3.2	508	2	Q5LJH8_MAGGR	Q5LJH8	magnaporthe
439	7	3.2	380	2	Q9A874_ORYSA	Q9A874	oryza sativ	512	7	3.2	510	2	Q4VMU6_9ROSI	Q4VMU6	securinega
440	7	3.2	381	2	Q5ZN27_PSEME	Q5ZN27	pseudomonas	513	7	3.2	510	2	Q63ZU9_MOUSE	Q63ZU9	mouse
441	7	3.2	382	2	Q5ZH39_PSEME	Q5ZH39	pseudomonas	514	7	3.2	515	1	C13A2_CASBL	C13A2	casabl
442	7	3.2	384	2	Q5ZN37_PSEPS	Q5ZN37	pseudomonas	515	7	3.2	515	2	Q4WY05_ASPFU	Q4WY05	aspeygiilus
443	7	3.2	386	2	Q54UN5_DICDI	Q54UN5	dicyostell	516	7	3.2	515	2	Q4RBH7_MOUSE	Q4RBH7	mouse
444	7	3.2	387	2	Q5ZH46_PSEPS	Q5ZH46	pseudomonas	517	7	3.2	520	2	Q21020_CABEL	Q21020	caenothabdi
445	7	3.2	393	2	Q50W31_ENTHI	Q50W31	entameoba h	518	7	3.2	523	2	Q8AV90_PETMA	Q8AV90	petromyxon
446	7	3.2	395	2	Q5AVI1_EMENT	Q5AVI1	entameoba h	519	7	3.2	528	2	Q6ZWM0_MOUSE	Q6ZWM0	mouse
447	7	3.2	396	2	Q537T8_NEIME	Q537T8	neisseria m	520	7	3.2	529	2	Q55SL4_CRYNE	Q55SL4	cryne
448	7	3.2	397	2	Q92D08_LISIN	Q92D08	liisteria in	521	7	3.2	529	2	Q5KH05_CRYNE	Q5KH05	cryptococcu
449	7	3.2	397	2	Q721P2_LISMF	Q721P2	liisteria mo	522	7	3.2	530	2	Q5BAZ9_CRYNE	Q5BAZ9	aspeygiilus
449	7	3.2	397	2	Q721P2_LISMF	Q721P2	liisteria mo	522	7	3.2	532	2	Q5BAZ9_CRYNE	Q5BAZ9	aspeygiilus
450	7	3.2	397	2	Q9RR46_LISMO	Q9RR46	liisteria mo	523	7	3.2	534	2	Q9HA36_HUMAN	Q9HA36	homo sapien
451	7	3.2	400	2	Q6C4S1_YARLI	Q6C4S1	yarrowia y	524	7	3.2	538	1	GLB1_EAST	GLB1	east
452	7	3.2	401	1	COAR_PAV	Q91720	paratocotia v	525	7	3.2	539	2	Q7S7G8_NEUCR	Q7S7G8	neutrospora
453	7	3.2	402	2	Q7S1Z5_ORYSA	Q7S1Z5	oryza sativ	526	7	3.2	539	2	Q8BJY0_MOUSE	Q8BJY0	m mouse
454	7	3.2	403	2	Q4K7X1_PSEPS	Q4K7X1	pseudomonas	527	7	3.2	540	2	Q6ALB3_DEEPS	Q6ALB3	desulfotale
455	7	3.2	406	2	Q4RUE6_TETNG	Q4RUE6	tetradon n	528	7	3.2	543	2	Q93S09_STRCO	Q93S09	strepomyce
456	7	3.2	410	2	PGK_SUTSO	Q4RUE6	tetradon n	529	7	3.2	544	2	Q9PE03_MOUSE	Q9PE03	mouse
457	7	3.2	410	2	Q71S47_ORYSA	Q71S47	oryza sativ	530	7	3.2	545	2	Q9HEK1_NEUCR	Q9HEK1	neutrospora
458	7	3.2	410	2	Q53NF8_ORYSA	Q53NF8	oryza sativ	531	7	3.2	547	2	Q4MYT8_TBRPA	Q4MYT8	thelateria p
459	7	3.2	411	2	Q13108_BRARE	Q13108	brachydantio	532	7	3.2	547	2	Q4Q8N0_LEIMA	Q4Q8N0	leishmania
460	7	3.2	411	2	Q57573_BRARE	Q57573	brachydantio	533	7	3.2	548	1	GNL3_XENTR	GNL3	xentropus tro
461	7	3.2	411	2	Q93369_BRARE	Q93369	brachydantio	534	7	3.2	548	2	Q9RW08_DEIRA	Q9RW08	delinococcus
462	7	3.2	412	1	APK1B_ARATH	P46573	arabidopsis	535	7	3.2	549	2	Q64B31_9ARCH	Q64B31	uncultured
463	7	3.2	412	1	Q7GAG3_ARATH	Q7GAG3	arabidopsis	536	7	3.2	549	2	Q8NAX3_HUMAN	Q8NAX3	homo sapien
464	7	3.2	413	1	RFIM_YEAST	P30775	saccharomyc	537	7	3.2	549	2	Q74F53_GEOSL	Q74F53	geobacter a
465	7	3.2	415	2	Q6RTJ1_CANGA	Q6RTJ1	candida gla	538	7	3.2	552	2	Q6A077_MOUSE	Q6A077	mouse
466	7	3.2	415	2	Q7XAM9_ORYSA	Q7XAM9	oryza sativ	539	7	3.2	555	2	Q81VP9_HUMAN	Q81VP9	homo sapien
467	7	3.2	416	2	Q7S116_NEUCR	Q7S116	neutrospora	540	7	3.2	555	2	Q9UDV4_LEIMA	Q9UDV4	leishmania
468	7	3.2	419	2	Q5B969_EMENT	Q5B969	aspeygiilus	541	7	3.2	557	2	Q4UB78_TREAN	Q4UB78	thelateria a
469	7	3.2	419	2	Q7VFD4_HELP	Q7VFD4	helicobacte	542	7	3.2	560	2	Q4NUZ0_9DELT	Q4NUZ0	anaeromyxob

543	7	3.2	561	2	Q4WL17 ASPFU	Q4WL17 aspergillus	616	7	3.2	730	2	Q7XL12 ORYSA	Q7XL12 oryza sativ
544	7	3.2	564	2	Q5BCX1 EMENTI	Q5BCX1 aspergillus	617	7	3.2	733	2	Q5L0T4 MAGGR	Q5L0T4 magnaporthe
545	7	3.2	566	2	Q7RW0 NEUCR	Q7RW0 neurospora	618	7	3.2	737	2	Q5BR1 EMENTI	Q5BR1 aspergillus
546	7	3.2	565	2	Q60T86 CABRR	Q60T86 caenorhabdi	619	7	3.2	738	2	Q5OTC4 ENTHI	Q5OTC4 entamoeba h
547	7	3.2	568	2	Q6NVMI XENTR	Q6NVMI xenopus tro	620	7	3.2	739	2	Q6VDS8 ORYSA	Q6VDS8 oryza sativ
548	7	3.2	574	2	Q5ZJ72 CHICK	Q5ZJ72 gallus gall	621	7	3.2	741	2	Q6LN22 CAEBR	Q6LN22 caenorhabdi
549	7	3.2	579	2	Q7TN27 RAT	Q7TN27 rattus norv	622	7	3.2	743	2	Q6ZN50 HUMAN	Q6ZN50 homo sapien
550	7	3.2	580	2	Q4FZ21 RAT	Q4FZ21 rattus norv	623	7	3.2	755	2	Q9X180 THEMA	Q9X180 thermocoga
551	7	3.2	581	2	Q8BX57 MOUSE	Q8BX57 mus musculu	624	7	3.2	768	2	Q8C585 MOUSE	Q8C585 mus musculu
552	7	3.2	582	2	Q91WB6 MOUSE	Q91WB6 mus musculu	625	7	3.2	771	2	Q8RUB5 METMA	Q8RUB5 methanosa
553	7	3.2	593	2	Q4Q28 LEIMA	Q4Q28 leishmania	626	7	3.2	771	2	Q7PXB1 ANOGA	Q7PXB1 anopheles g
554	7	3.2	594	2	Q81129 CAEBL	Q81129 caenorhabdi	627	7	3.2	784	1	QBP PLAFG	QBP PLAFG
555	7	3.2	595	1	KINB_PSEAE	Q34206 pseudomonas	628	7	3.2	785	2	Q7D3D9 AGRT5	Q7D3D9 agrobacteri
556	7	3.2	595	2	Q9HT87 PSEAE	Q9HT87 pseudomonas	629	7	3.2	789	2	Q5RB27 PONPY	Q5RB27 pongo pygma
557	7	3.2	597	2	Q4WV4 ASPFU	Q4WV4 aspergillus	630	7	3.2	798	2	Q5BR22 EMENTI	Q5BR22 aspergillus
558	7	3.2	597	2	Q7U2T5 MYCBO	Q7U2T5 mycobacteri	631	7	3.2	798	2	Q9NVA7 HUMAN	Q9NVA7 homo sapien
559	7	3.2	600	2	Q4J3V8 AZOVI	Q4J3V8 azotobacter	632	7	3.2	799	2	Q7Z3K9 HUMAN	Q7Z3K9 homo sapien
560	7	3.2	602	2	Q9P8C3 9HYPO	Q9P8C3 acetonitium	633	7	3.2	799	2	Q8V3M3 SWPV	Q8V3M3 swinepox vi
561	7	3.2	603	2	Q9NMW6 HUMAN	Q9NMW6 homo sapien	634	7	3.2	801	2	Q8V3M3 SWPV	Q8V3M3 swinepox vi
562	7	3.2	606	2	Q9Z2J3 RAT	Q9Z2J3 rattus norv	635	7	3.2	804	2	Q6P1Z0 HUMAN	Q6P1Z0 homo sapien
563	7	3.2	607	2	Q7TWM1 MOUSE	Q7TWM1 mus musculu	636	7	3.2	805	2	Q5EBM3 HUMAN	Q5EBM3 homo sapien
564	7	3.2	610	2	Q4HZK6 GIBZE	Q4HZK6 gibberella	637	7	3.2	807	2	Q7T1B6 FUGRU	Q7T1B6 fugu rubrip
565	7	3.2	612	2	Q7X678 CHLRE	Q7X678 chlamydomon	638	7	3.2	807	2	Q8AXT1 FUGRU	Q8AXT1 xyella fas
566	7	3.2	614	2	Q55YK6 CRYNE	Q55YK6 cryptococcu	639	7	3.2	809	1	PBPA XYLPT	PBPA XYLPT
567	7	3.2	614	2	Q5YKID2 CRYNE	Q5YKID2 cryptococcu	640	7	3.2	809	1	PBPA XYLPT	PBPA XYLPT
568	7	3.2	615	2	Q6ZWR5 MOUSE	Q6ZWR5 mus musculu	641	7	3.2	809	2	Q8CGG1 MOUSE	Q8CGG1 mus musculu
569	7	3.2	619	2	Q4KKX0 HUMAN	Q4KKX0 homo sapien	642	7	3.2	809	2	Q6PAR2 MOUSE	Q6PAR2 mus musculu
570	7	3.2	621	2	Q5BDB4 EMENTI	Q5BDB4 aspergillus	643	7	3.2	812	2	Q7R118 GIATA	Q7R118 giardia lam
571	7	3.2	621	2	Q4WQLE ASPFU	Q4WQLE aspergillus	644	7	3.2	816	2	Q5ZL41 CHICK	Q5ZL41 gallus gall
572	7	3.2	621	2	Q49047 TOBAC	Q49047 nicotiana t	645	7	3.2	819	2	Q8UK13 AGRT5	Q8UK13 agrobacteri
573	7	3.2	624	1	PARN MOUSE	Q8V6G3 mus musculu	646	7	3.2	822	2	Q91288 PLEMA	Q91288 plautodelas
574	7	3.2	630	2	Q5P670 AZOSE	Q5P670 azotarcus sp	647	7	3.2	824	1	GBP PLAF7	GBP PLAF7
575	7	3.2	633	1	P6GA HUMAN	Q86P4 homo sapien	648	7	3.2	828	1	SKOR ARATH	SKOR ARATH
576	7	3.2	633	2	Q98LE9 RHILLO	Q98LE9 rhizobium l	649	7	3.2	831	2	Q7S5Z4 NEUCR	Q7S5Z4 neurospora
577	7	3.2	633	2	Q926H1 RHIME	Q926H1 rhizobium m	650	7	3.2	842	2	Q54MH2 DICDI	Q54MH2 dicystoteli
578	7	3.2	635	1	SYT CHILAB	Q514C1 chlamydomon	651	7	3.2	843	2	Q6AXC9 MOUSE	Q6AXC9 mus musculu
579	7	3.2	637	2	Q5DZP9 VIBF1	Q5DZP9 vibrio fibc	652	7	3.2	849	2	Q9XR68 DROME	Q9XR68 drosophila
580	7	3.2	639	1	PARN HUMAN	Q95453 homo sapien	653	7	3.2	856	2	Q8H1P8 ARATH	Q8H1P8 arabidopsis
581	7	3.2	639	1	PARN_PONPY	Q5RC51 pongo pygma	654	7	3.2	857	1	GELA DICDI	GELA DICDI
582	7	3.2	651	2	Q6PFR6 BRARE	Q6PFR6 brachydanio	655	7	3.2	857	2	Q5C12 DICDI	Q5C12 dicystoteli
583	7	3.2	656	1	R1PK1 MOUSE	Q60855 mus musculu	656	7	3.2	861	2	Q9VLX2 DROME	Q9VLX2 drosophila
584	7	3.2	656	2	Q9Z2J4 RAT	Q9Z2J4 rattus norv	657	7	3.2	871	2	Q9VP66 THEYO	Q9VP66 thermocoga
585	7	3.2	669	2	Q4J0E8 AZOVI	Q4J0E8 azotobacter	658	7	3.2	871	2	Q8RBE6 THETN	Q8RBE6 thermocoga
586	7	3.2	669	2	Q88CB7 PSEBK	Q88CB7 pseudomonas	659	7	3.2	875	2	Q5Y2P9 DERPT	Q5Y2P9 dermatophag
587	7	3.2	669	2	Q9HT08 PSEAE	Q9HT08 pseudomonas	660	7	3.2	876	1	MISP SARSC	MISP SARSC
588	7	3.2	676	2	Q84Q65 ORYSA	Q84Q65 oryza sativ	661	7	3.2	876	2	Q17951 CAEBL	Q17951 caenorhabdi
589	7	3.2	687	2	Q5P6M5 AZOSE	Q5P6M5 azotarcus sp	662	7	3.2	876	2	Q7TN28 MOUSE	Q7TN28 mus musculu
590	7	3.2	688	1	DNJL1 THEMA	Q9W4V5 thermocoga	663	7	3.2	877	1	INCE CHICK	INCE CHICK
591	7	3.2	690	1	EPB42 HUMAN	P16452 homo sapien	664	7	3.2	878	1	Q5BP2 HUMAN	Q5BP2 homo sapien
592	7	3.2	691	2	Q4YB97 HUMAN	Q4YB97 homo sapien	665	7	3.2	878	2	Q5AS58 EMENTI	Q5AS58 aspergillus
593	7	3.2	692	1	MISP DERPA	Q96720 dermatophag	666	7	3.2	880	1	INCE MOUSE	INCE MOUSE
594	7	3.2	692	1	Q7SE14 NEUCR	Q7SE14 neurospora	667	7	3.2	880	2	Q8PGY9 XANAC	Q8PGY9 xanthomonas
595	7	3.2	695	2	Q521Y5 MAGGR	Q521Y5 magnaporthe	668	7	3.2	881	2	Q5H3V2 XANOR	Q5H3V2 xanthomonas
596	7	3.2	695	2	Q7ZUM8 BRARE	Q7ZUM8 brachydanio	669	7	3.2	886	2	Q9VZ04 DROME	Q9VZ04 drosophila
597	7	3.2	698	2	Q8NC60 HUMAN	Q8NC60 homo sapien	670	7	3.2	892	2	Q4SC70 TETNG	Q4SC70 tetraodon n
598	7	3.2	698	2	Q8BS08 HUMAN	Q8BS08 homo sapien	671	7	3.2	899	2	Q7OE23 NOGCA	Q7OE23 anopheles g
599	7	3.2	700	1	SYNG RAT	Q9YJCK9 rattus norv	672	7	3.2	908	2	Q5ON06 MOUSE	Q5ON06 mus musculu
600	7	3.2	701	2	Q4K3R2 PSEBP	Q4K3R2 pseudomonas	673	7	3.2	914	2	Q50UG3 ENTHI	Q50UG3 entamoeba h
601	7	3.2	701	2	Q9HTM0 PSEAB	Q9HTM0 pseudomonas	674	7	3.2	917	2	Q5THY6 HUMAN	Q5THY6 homo sapien
602	7	3.2	702	2	Q88BE4 PSEBK	Q88BE4 pseudomonas	675	7	3.2	918	2	Q5Y192 HUMAN	Q5Y192 homo sapien
603	7	3.2	702	2	Q88CB1 PSEBK	Q88CB1 pseudomonas	676	7	3.2	919	1	INCE HUMAN	INCE HUMAN
604	7	3.2	704	2	Q4P665 USITMA	Q4P665 usellago ma	677	7	3.2	922	2	Q8RFX5 ARATH	Q8RFX5 arabidopsis
605	7	3.2	705	2	Q82496 ARATH	Q82496 arabidopsis	678	7	3.2	923	2	Q60424 HUMAN	Q60424 homo sapien
606	7	3.2	707	2	Q4Z2Z0 PSEBY	Q4Z2Z0 pseudomonas	679	7	3.2	937	2	Q7ZUV4 BRARE	Q7ZUV4 brachydanio
607	7	3.2	708	2	Q873A0 NEUCR	Q873A0 neurospora	680	7	3.2	940	2	Q6CAR0 YARLI	Q6CAR0 yarrowia li
608	7	3.2	712	2	Q4K686 PSEF5	Q4K686 pseudomonas	681	7	3.2	945	2	Q7P6J0 FUSNV	Q7P6J0 fusobacteri
609	7	3.2	714	1	BEGL MYCTU	Q07170 mycobacteri	682	7	3.2	955	2	Q6BK79 DEBHA	Q6BK79 debaryomyce
610	7	3.2	718	2	Q5LI14 MAGGR	Q5LI14 magnaporthe	683	7	3.2	958	2	Q7SLD7 ASHGO	Q7SLD7 ashbya gos
611	7	3.2	718	2	Q9NVQ0 HUMAN	Q9NVQ0 homo sapien	684	7	3.2	960	2	Q8L671 ORYSA	Q8L671 oryza sativ
612	7	3.2	720	2	Q73U43 MYCPA	Q73U43 mycobacteri	685	7	3.2	964	2	Q42489 CHICK	Q42489 gallus gall
613	7	3.2	728	2	Q7SELO NEUCR	Q7SELO neurospora	686	7	3.2	967	2	Q84FK8 ENTAG	Q84FK8 enterobacte
614	7	3.2	728	2	Q7ZJ86 THEIT2	Q7ZJ86 thermus the	687	7	3.2	968	2	Q990M5 YVIRU	Q990M5 infectious
615	7	3.2	728	2	Q5SIV7 THEIT8	Q5SIV7 thermus the	688	7	3.2	968	2	Q990M6 YVIRU	Q990M6 infectious

689	7	3.2	968	2	Q990M7_YVIRU	Q990m7_infectious	762	7	3.2	1254	2	Q29122_PIG	Q29122_bus_acrofa
690	7	3.2	968	2	Q990M8_YVIRU	Q990m8_infectious	763	7	3.2	1255	2	Q6B335_BRARE	Q6B335_brachydantio
691	7	3.2	968	2	Q990P3_YVIRU	Q990p3_infectious	764	7	3.2	1256	2	Q4WJW2_ASPFU	Q4WJW2_aspergillus
692	7	3.2	968	2	Q990Q1_YVIRU	Q990q1_infectious	765	7	3.2	1257	1	FLII_CAEEL	P34266_caenorhabdi
693	7	3.2	970	2	Q4TLU5_95PHN	Q4tlus_elytrioabact	766	7	3.2	1262	1	OSTEM7_HUMAN	OSTEM7_homo_sapien
694	7	3.2	972	1	POLS_IPNVJ	P05844_infectious	767	7	3.2	1265	1	MYO6_MOUSE	Q64331_mus_musculu
695	7	3.2	972	2	Q7TLC2_YVIRU	Q7tlc2_infectious	768	7	3.2	1267	2	Q9N8F0_STRPU	Q9N8F0_strongyloce
696	7	3.2	972	2	Q827J3_YVIRU	Q827j3_infectious	769	7	3.2	1267	2	Q68R03_BRARS	Q68R03_brachydantio
697	7	3.2	972	2	Q8TK05_YVIRU	Q8tk05_infectious	770	7	3.2	1268	2	Q7RG05_PLAYO	Q7RG05_plasmodium
698	7	3.2	972	2	Q9WLB3_YVIRU	Q9wlb3_infectious	771	7	3.2	1285	1	Q91BD1_CHICK	Q91BD1_gallus_gall
699	7	3.2	972	2	Q60E16_ORYSA	Q60e16_oryza_sativ	772	7	3.2	1285	1	MYO6_HUMAN	Q9um54_homo_sapien
700	7	3.2	977	2	Q60E16_ORYSA	Q60e16_oryza_sativ	773	7	3.2	1285	2	OSTEM6_HUMAN	OSTEM6_homo_sapien
701	7	3.2	980	2	Q5U7X8_YVIRU	Q5u7x8_infectious	774	7	3.2	1291	2	Q6Y7W8_MOUSE	Q6Y7W8_mus_musculu
702	7	3.2	980	2	Q5U7Y4_YVIRU	Q5u7y4_infectious	775	7	3.2	1292	2	Q68R02_BRARS	Q68R02_brachydantio
703	7	3.2	980	2	Q5U7Y6_YVIRU	Q5u7y6_infectious	776	7	3.2	1298	2	Q72312_HUMAN	Q72312_homo_sapien
704	7	3.2	980	2	Q5U7Z0_YVIRU	Q5u7z0_infectious	777	7	3.2	1299	2	Q6Y7W6_HUMAN	Q6Y7W6_homo_sapien
705	7	3.2	982	2	Q4KZV0_TETNG	Q4kzv0_tetradodon n	778	7	3.2	1299	2	Q6GLR7_XENLA	Q6GLR7_xenopus_lae
706	7	3.2	982	2	Q56YV1_ARATH	Q56yv1_arabidopsis	779	7	3.2	1299	1	STNG_MOUSE	Q56lv8_mus_musculu
707	7	3.2	991	2	Q56YV2_ARATH	Q56yv2_arabidopsis	780	7	3.2	1306	1	STNG_HUMAN	Q5um22_homo_sapien
708	7	3.2	991	2	Q56YR0_ARATH	Q56yr0_arabidopsis	781	7	3.2	1314	2	Q9UP78_HUMAN	Q9up78_homo_sapien
709	7	3.2	997	2	Q50VLA_ENTHI	Q50v14_entamoeba h	782	7	3.2	1315	2	Q75137_HUMAN	Q75137_homo_sapien
710	7	3.2	997	2	Q54YB5_DICDI	Q54ybs_dicystosteli	783	7	3.2	1329	2	Q4KME6_BRARE	Q4KME6_brachydantio
711	7	3.2	1002	1	CLNM_HUMAN	Q54ybs_dicystosteli	784	7	3.2	1335	1	SRC3_YEAST	P33332_saccharomyc
712	7	3.2	1002	2	Q6NUQ2_HUMAN	Q6nuq2_homo_sapien	785	7	3.2	1336	2	Q6N8E5_NEUCR	Q6n8e5_neurospora
713	7	3.2	1018	2	Q753P3_ASHGO	Q753p3_asbhya_goss	786	7	3.2	1367	2	Q4LB49_HUMAN	Q4lb49_homo_sapien
714	7	3.2	1018	2	Q5VS68_ORYSA	Q5vs68_oryza_sativ	787	7	3.2	1374	2	Q81586_PLAF7	Q81586_plasmodium
715	7	3.2	1028	2	Q5VU16_HUMAN	Q5vu16_homo_sapien	788	7	3.2	1377	2	TRHY_RABIT	P37709_cryptosleagus
716	7	3.2	1028	2	Q04604_ARATH	Q04604_arabidopsis	789	7	3.2	1407	1	Q4T3R1_TETNG	Q4t3r1_tetradodon n
717	7	3.2	1032	2	Q4XVNZ_PLACH	Q4xvnz_plasmodium	790	7	3.2	1414	2	Q5R129_BRARE	Q5r129_brachydantio
718	7	3.2	1032	2	Q6QM03_AEGTA	Q6qm03_aegilops ta	791	7	3.2	1418	2	Q5VU15_HUMAN	Q5vu15_homo_sapien
719	7	3.2	1036	2	Q4WYU6_ASPFU	Q4wyu6_aspergillus	792	7	3.2	1426	2	Q8K3V7_MOUSE	Q8k3v7_mus_musculu
720	7	3.2	1042	2	Q4Q5S2_LEIMA	Q4q5s2_leishmania	793	7	3.2	1431	2	Q4N018_THERA	Q4n018_thelazia p
721	7	3.2	1045	2	Q9NXY1_DROME	Q9nxy1_drosophila	794	7	3.2	1433	2	Q4YZ38_PLABE	Q4yz38_plasmodium
722	7	3.2	1046	2	Q6NN85_DROME	Q6nn85_drosophila	795	7	3.2	1440	2	Q53NR4_ORYSA	Q53nr4_oryza_sativ
723	7	3.2	1048	2	Q80TV1_MOUSE	Q80tv1_mus_musculu	796	7	3.2	1488	2	Q4WC64_ASPFU	Q4wc64_aspergillus
724	7	3.2	1052	1	CLNM_MOUSE	Q80tv1_mus_musculu	797	7	3.2	1497	2	Q4RKB4_TETNG	Q4rkb4_tetradodon n
725	7	3.2	1074	2	Q82491_ARATH	Q82491_arabidopsis	798	7	3.2	1521	2	Q7R2Y4_GIALA	Q7r2y4_giardia lam
726	7	3.2	1079	2	Q5R844_BRARE	Q5r844_brachydantio	799	7	3.2	1539	2	Q7R056_GIALA	Q7r056_giardia lam
727	7	3.2	1089	2	Q75B21_ASHGO	Q75b21_asbhya_goss	800	7	3.2	1552	2	Q519G4_ENTHI	Q519g4_entamoeba h
728	7	3.2	1091	2	Q8DPR7_STRR6	Q8dpr7_streptococc	801	7	3.2	1575	2	Q54R77_DICDI	Q54r77_dicystosteli
729	7	3.2	1091	2	Q970Q0_STRPN	Q97q0_streptococc	802	7	3.2	1603	2	Q7S6J7_NEUCR	Q7s6j7_neurospora
730	7	3.2	1096	2	Q4SP07_TETNG	Q4sp07_tetradodon n	803	7	3.2	1617	2	Q527G3_MAGCR	Q527g3_magnaporthe
731	7	3.2	1099	2	Q86ZJ5_PODAN	Q86zj5_podospora a	804	7	3.2	1620	2	Q5H962_HUMAN	Q5h962_homo_sapien
732	7	3.2	1100	2	Q9NV82_HUMAN	Q9nv82_homo_sapien	805	7	3.2	1672	2	Q9Y1T1_TOXGO	Q9y1t1_toxoplasma
733	7	3.2	1106	1	ITAT_RAT	Q63258_rattus norv	806	7	3.2	1690	2	Q870V1_NEUCR	Q870v1_neurospora
734	7	3.2	1109	2	Q54TT8_DICDI	Q54tt8_dicystosteli	807	7	3.2	1771	2	Q5DPT4_MOUSE	Q5dpt4_mus_musculu
735	7	3.2	1135	2	Q5HXZ9_RAT	Q5hxz9_rattus norv	808	7	3.2	1796	2	Q4SUA7_TETNG	Q4sua7_tetradodon n
736	7	3.2	1138	2	Q4SKZ1_TETNG	Q4skz1_tetradodon n	809	7	3.2	1830	2	TRHY_HUMAN	Q07283_homo_sapien
737	7	3.2	1146	2	Q8CH18_MOUSE	Q8ch18_mus_musculu	810	7	3.2	1898	1	Q925Q1_MOUSE	Q925q1_mus_musculu
738	7	3.2	1146	2	Q4RYK6_TETNG	Q4ryk6_tetradodon n	811	7	3.2	1902	2	Q640Q1_MOUSE	Q640q1_mus_musculu
739	7	3.2	1150	2	Q4RYK5_HUMAN	Q4ryk5_homo_sapien	812	7	3.2	1902	2	Q5VU13_HUMAN	Q5vu13_homo_sapien
740	7	3.2	1150	2	Q81X12_HUMAN	Q81x12_homo_sapien	813	7	3.2	1943	2	Q88H71_PSEPK	Q88h71_pseudomonas
741	7	3.2	1150	2	Q5VUP6_HUMAN	Q5vup6_homo_sapien	814	7	3.2	1970	2	Q601N7_MYCHY	Q601n7_mycoplasma
742	7	3.2	1154	2	Q55C65_DICDI	Q55c65_dicystosteli	815	7	3.2	2047	2	Q4REJ8_TETNG	Q4rej8_tetradodon n
743	7	3.2	1157	2	Q641G3_XENLA	Q641g3_xenopus lae	816	7	3.2	2075	1	BAT2_MOUSE	Q7ezc1_mus_musculu
744	7	3.2	1166	2	Q4IO90_GIBZE	Q4iog0_gibberella	817	7	3.2	2158	1	Q5U246_XENLA	Q5u246_xenopus lae
745	7	3.2	1166	2	Q96Z19_PLAF7	Q96z19_plasmodium	818	7	3.2	2161	1	BAT2_RAT	Q14497_h-at-rich i
746	7	3.2	1178	2	Q6AHQ1_CAEEL	Q6ahq1_caenorhabdi	819	7	3.2	2285	1	AR1A_HUMAN	Q5uz48_homo_sapien
747	7	3.2	1179	1	ITAT_MOUSE	Q61738_mus_musculu	820	7	3.2	2428	2	Q80ZV8_MOUSE	Q80zv8_mus_musculu
748	7	3.2	1187	2	Q8BWG4_MOUSE	Q8bwg4 m mus muscu	821	7	3.2	2429	2	CBP_MOUSE	Q56lv8_mus_musculu
749	7	3.2	1192	2	Q7KS05_DROME	Q7ks05_drosophila	822	7	3.2	2441	1	CBP_HUMAN	P45f81_mus_musculu
750	7	3.2	1193	2	Q8IMU8_DROME	Q8imu8_drosophila	823	7	3.2	2442	1	Q6UH9_RAT	Q6uh9_homo_sapien
751	7	3.2	1193	2	Q8KJ29_RHILU	Q8kj29_rhizobium l	824	7	3.2	2442	2	Q4LEJ8_HUMAN	Q4lej8_homo_sapien
752	7	3.2	1207	2	Q4P3H6_USTMA	Q4p3h6_ustilago ma	825	7	3.2	2472	2	Q4S8B2_TETNG	Q4s8b2_tetradodon n
753	7	3.2	1208	2	Q96T49_HUMAN	Q96t49_homo_sapien	826	7	3.2	2473	2	Q7WRU0_9NOST	Q7wrj0_anabaena cl
754	7	3.2	1216	2	Q4S2D6_TETNG	Q4s2d6_tetradodon n	827	7	3.2	2635	2	Q60WB1_CABER	Q60wb1_caenorhabdi
755	7	3.2	1219	2	Q523G6_MAGCR	Q523g6_magnaporthe	828	7	3.2	2696	2	Q7S935_NEUCR	Q7s935_neurospora
756	7	3.2	1224	2	Q5BKU5_HUMAN	Q5bku5_homo_sapien	829	7	3.2	2826	2	Q93326_CAEEL	Q93326_caenorhabdi
757	7	3.2	1227	2	Q5U2J6_XENLA	Q5u2j6_xenopus lae	830	7	3.2	2862	2	Q5H924_HUMAN	Q5h924_homo_sapien
758	7	3.2	1239	2	Q627P2_CABER	Q627p2_caenorhabdi	831	7	3.2	3407	2	Q7OKJ6_BACAM	Q7okj6_bacillus am
759	7	3.2	1251	2	OSTEM5_HUMAN	OSTEM5_homo_sapien	832	7	3.2	3586	2	Q4PEL1_USTMA	Q4pel1_ustilago ma
760	7	3.2	1253	2	Q6B334_BRARE	Q6b334_brachydantio	833	7	3.2	3893	2	Q9BK91_STRPU	Q9bk91_strongyloce
761	7	3.2	1253	2			834	7	3.2	3908	2		

835	7	3.2	4037	2	Q4WPX3_ASEFU	Q4WPX3 aspergillus	908	6	2.7	66	2	Q6R5D6_MOUSE	Q6r5d6 mus musculus
836	7	3.2	4374	1	HUWE1_HUMAN	Q7a6z7 homo sapien	909	6	2.7	67	1	SLYX_RHITO	Q9g6w1 rhizobium 1
837	7	3.2	4374	2	Q4G2Z2_HUMAN	Q4g2z2 homo sapien	910	6	2.7	67	2	Q97X55_SULISO	Q97x55 sulfolobus
838	7	3.2	4374	2	Q4G2Z1_MOUSE	Q4g2z1 mus musculus	911	6	2.7	67	2	Q7S1J2_NEUCR	Q7s1j2 neurospora
839	7	3.2	4377	1	HUWE1_MOUSE	Q7luy8 mus musculus	912	6	2.7	67	2	Q6C3N6_YARLI	Q6c3n6 yarrowia 1i
840	7	3.2	4378	2	Q4UG03_MOUSE	Q4j903 mus musculus	913	6	2.7	67	2	Q9VCP3_DROME	Q9vcp3 drosophila
841	7	3.2	4473	1	PLEC1_CRIGR	Q9j155 cricetulus	914	6	2.7	67	2	Q8GVJ8_9POAL	Q8gvj8 agrostis ac
842	7	3.2	4684	1	PLEC1_HUMAN	Q15149 homo sapien	915	6	2.7	67	2	Q8IX90_BACAN	Q8ix90 bacillus an
843	7	3.2	4687	1	PLEC1_RAT	P30427 rattus norv	916	6	2.7	67	2	Q6MBB3_PARUV	Q6mbb3 paracaulamd
844	7	3.2	4796	2	Q4Q363_LEIMA	Q4q363 leishmania	917	6	2.7	67	2	Q9MLQ7_SHIVI	Q9mlq7 human immun
845	7	3.2	6885	1	SYNE2_HUMAN	Q8w4h0 homo sapien	918	6	2.7	68	2	Q5V4F7_HALMA	Q5v4f7 haloarcula
846	7	3.2	8348	2	Q4S6Y5_TETNG	Q4s6y5 tetraodon n	919	6	2.7	68	2	Q7UD59_SHIFL	Q7ud59 shigella fl
847	7	3.2	8407	2	Q7RTM4_HUMAN	Q7rtm4 homo sapien	920	6	2.7	69	2	Q5I2P0_CONLE	Q5i2p0 conus leopa
848	7	3.2	8654	2	OSTY05_BRARE	OSTy05 brachydario	921	6	2.7	69	2	Q6EP69_ORYSA	Q6ep69 oryza sativ
849	7	3.2	8797	1	SYNE1_HUMAN	Q8tf91 homo sapien	922	6	2.7	69	2	Q8RTM2_CORER	Q8rtm2 corynabace
850	6	2.7	28	2	Q7RF66_PLAYO	Q7rf66 plasmodium	923	6	2.7	69	2	Q8QNE9_9PHYC	Q8qne9 ectorcarpac
851	6	2.7	30	2	Q7RF64_PLAYO	Q7rf64 plasmodium	924	6	2.7	70	2	Q84CV8_9PACT	Q84cv8 uncultured
852	6	2.7	38	2	Q16493_HUMAN	Q16493 homo sapien	925	6	2.7	70	2	Q6NHG3_CORDI	Q6nhg3 cotrynebacte
853	6	2.7	30	2	Q4RAC9_TETNG	Q4rac9 tetraodon n	926	6	2.7	70	2	Q6UG85_YVIRU	Q6ug85 sulfolobus
854	6	2.7	34	2	Q6B113_YEAST	Q6b113 saccharomyc	927	6	2.7	70	2	Q8BCI8_NPVEM	Q8bc18 bombyx mori
855	6	2.7	34	2	Q7RI25_PLAYO	Q7ri25 plasmodium	928	6	2.7	71	2	Q8UP56_AGRTS	Q8uf56 agrobacteri
856	6	2.7	36	2	Q5SJU3_THERT	Q5sj13 thermus the	929	6	2.7	71	2	Q8UP56_AGRTS	Q8uf56 agrobacteri
857	6	2.7	37	1	Q6D28_LUPAN	P09930 lupinus ang	930	6	2.7	72	2	Q7V8B2_PROMM	Q7v8b2 prochlororo
858	6	2.7	37	2	Q41794_SHIVI	Q41794 human immun	931	6	2.7	72	2	Q8LCW7_ARATH	Q8lcw7 arabidopsis
859	6	2.7	37	2	Q89598_SHIVI	Q89598 human immun	932	6	2.7	72	2	Q94A69_ARATH	Q94a69 arabidopsis
860	6	2.7	37	2	Q8UTU3_SHIVI	Q8ut3 human immun	933	6	2.7	72	2	Q6HR26_BACAN	Q6hr26 bacillus an
861	6	2.7	37	2	Q89489_SHIVI	Q89489 human immun	934	6	2.7	72	2	Q4SUW2_TETNG	Q4suw2 tetraodon n
862	6	2.7	37	2	Q8PXM6_SHIVI	Q8pxm6 human immun	935	6	2.7	73	2	Q9DUJ2_SHIVI	Q9duj2 human immun
863	6	2.7	37	2	Q8PXM7_SHIVI	Q8pxm7 human immun	936	6	2.7	73	2	Q8TLZ6_METAC	Q8tlz6 methanosarc
864	6	2.7	39	2	Q9GCF1_9GAMA	Q9gcf1 human herpe	937	6	2.7	73	2	Q8NEB9_HUMAN	Q8neb9 homo sapien
865	6	2.7	40	2	Q6K180_MYCMO	Q6k180 mycoplasma	938	6	2.7	73	2	Q6LCK2_MOUSE	Q6lck2 mus musculu
866	6	2.7	41	2	Q5I3T1_SHIVI	Q5i3t1 human immun	939	6	2.7	74	2	Q8AYX6_9GAMA	Q8ayx6 porcine lym
867	6	2.7	42	1	BDJ3_BOVIN	P46171 bos taurus	940	6	2.7	74	2	Q6B8C2_9ACAR	Q6b8c2 iodice paci
868	6	2.7	44	2	Q7JKA9_TREDE	Q7jka9 treponema d	941	6	2.7	74	2	Q4YR44_PLABE	Q4yfr4 plasmodium
869	6	2.7	45	1	SEPU_BACSU	Q7wy58 bacillus su	942	6	2.7	75	2	Q70B17_HDV	Q70b17 hepatitis d
870	6	2.7	45	2	Q5C4D6_SCHJA	Q5c4d6 schistosoma	943	6	2.7	75	2	Q8U397_PYRPU	Q8u397 pyrococcus
871	6	2.7	45	2	Q7RGN2_PLAYO	Q7r92 plasmodium	944	6	2.7	75	2	Q6CB05_YARLI	Q6cb05 yarrowia 1i
872	6	2.7	48	2	Q9ZGZ0_YERPE	Q9zgz0 yersinia pe	945	6	2.7	76	2	Q7XIJ7_ORYSA	Q7xij7 oryza sativ
873	6	2.7	48	2	Q74Y68_YERPE	Q74y68 yersinia pe	946	6	2.7	76	2	Q8L9X5_ARATH	Q8l9x5 arabidopsis
874	6	2.7	48	2	Q8E517_STRAS	Q8e517 streptococc	947	6	2.7	76	1	Q83CM2_COXBU	Q83cm2 coxiella bu
875	6	2.7	49	2	Q6COK4_YARLI	Q6cok4 yarrowia 1i	948	6	2.7	78	1	Q83CM2_COXBU	Q83cm2 coxiella bu
876	6	2.7	49	2	Q605P1_METCA	Q605p1 methylococc	949	6	2.7	78	1	Q83CM2_COXBU	Q83cm2 coxiella bu
877	6	2.7	51	2	Q8TABO_HUMAN	Q8tabo homo sapien	950	6	2.7	78	1	Q83CM2_COXBU	Q83cm2 coxiella bu
878	6	2.7	51	2	Q5JRR7_HUMAN	Q5jrr7 homo sapien	951	6	2.7	78	1	Q83CM2_COXBU	Q83cm2 coxiella bu
879	6	2.7	51	2	Q4YMG0_PLABE	Q4ymg0 plasmodium	952	6	2.7	78	1	Q83CM2_COXBU	Q83cm2 coxiella bu
880	6	2.7	52	2	Q5MEO6_BACSK	Q5mew6 bacillus cl	953	6	2.7	78	2	Q6EG44_SHIVI	Q6eg44 shiv1
881	6	2.7	52	2	Q5MEO8_BACSK	Q5mew8 bacillus cl	954	6	2.7	78	2	Q6EG53_SHIVI	Q6eg53 shiv1
882	6	2.7	52	2	Q9ZCD5_LISIN	Q9zcd5 listeria in	955	6	2.7	78	2	Q77YF9_SHIVI	Q77yf9 human immun
883	6	2.7	53	1	Q804A3_PSEAM	Q804a3 pseudopleur	956	6	2.7	78	2	Q77YF9_SHIVI	Q77yf9 human immun
884	6	2.7	53	1	BDCT_BOVIN	Q18815 bos taurus	957	6	2.7	79	2	Q6LIZ3_DROME	Q6liz3 drosophila
885	6	2.7	53	2	Q8PF63_LEPIN	Q8pf63 leptospira	958	6	2.7	79	2	Q5OWV6_ENTHI	Q5owv6 entamoeba h
886	6	2.7	53	2	Q8PBM7_ECOL6	Q8pbm7 escherichia	959	6	2.7	79	2	Q9ACB0_BRUAB	Q9acb0 brucella ab
887	6	2.7	54	2	Q16031_HUMAN	Q16031 homo sapien	960	6	2.7	79	2	Q5GS73_WOLTE	Q5gs73 wolbachia s
888	6	2.7	55	2	Q6UG86_YVIRU	Q6ug86 sulfolobus	961	6	2.7	79	2	Q4T045_TETNG	Q4t045 tetraodon n
889	6	2.7	56	2	Q90175_SHIVI	Q90175 human immun	962	6	2.7	80	2	Q656V6_ORYSA	Q656v6 oryza sativ
890	6	2.7	57	2	Q4M8Z3_ASPFU	Q4m8z2 aspergillus	963	6	2.7	80	2	Q8LQ79_ORYSA	Q8lq79 oryza sativ
891	6	2.7	58	2	Q26585_METTH	Q26585 methanobact	964	6	2.7	80	2	Q9KMZ5_XANOR	Q9kmz5 xanthomonas
892	6	2.7	58	2	Q4UVI0_CORUS	Q4uvi0 corynebacte	965	6	2.7	80	2	Q6ZUJ4_BACLD	Q6zu4 bacillus li
893	6	2.7	59	2	Q691T7_ORYSA	Q691t7 oryza sativ	966	6	2.7	82	2	Q4US45_XANCP	Q4uas5 xanthomonas
894	6	2.7	59	2	Q6TRU3_YVIRU	Q6tru3 sulfolobus	967	6	2.7	81	2	Q8PBA1_XANCP	Q8pba1 xanthomonas
895	6	2.7	60	2	Q7IUA5_BOVIN	Q7luas bos taurus	968	6	2.7	81	2	Q9YEB1_AERPE	Q9yeb1 aeropyrum p
896	6	2.7	60	2	Q5M5G7_BOVIN	Q5m5g7 bos taurus	969	6	2.7	81	2	Q6IHR2_DROME	Q6ih2 drosophila
897	6	2.7	63	2	Q68HK2_XANOR	Q68hk2 xanthomonas	970	6	2.7	82	2	Q93KQ2_YEREN	Q93kq2 yersinia en
898	6	2.7	63	2	Q63NTR_BURPS	Q63ntr burkholderi	971	6	2.7	82	2	Q6EPV5_ORYSA	Q6epv5 oryza sativ
899	6	2.7	65	2	Q6G9P1_STAAR	Q6g9p1 staphylococ	972	6	2.7	82	2	Q6GVH8_FRANF	Q6gvh8 fragarita an
900	6	2.7	65	2	Q6H89_STAAR	Q6h89 staphylococ	973	6	2.7	82	2	Q8VH78_MUSSA	Q8vh78 mus saxicol
901	6	2.7	65	2	Q7VA99_PROMA	Q7va99 prochlororo	974	6	2.7	83	2	Q8VH78_MUSSA	Q8vh78 mus saxicol
902	6	2.7	65	2	Q5HG44_STAAC	Q5hg44 staphylococ	975	6	2.7	83	2	Q70ZV6_MUSPA	Q70zv6 anopheles g
903	6	2.7	65	2	Q7A100_STAAM	Q7a100 staphylococ	976	6	2.7	83	2	Q6K4B4_ORYSA	Q6k4b4 oryza sativ
904	6	2.7	65	2	Q7ASUC_STAAM	Q7asuc staphylococ	977	6	2.7	84	2	Q5ECA3_SHIVI	Q5eca3 human immun
905	6	2.7	65	2	Q99UP9_STAAM	Q99uf9 staphylococ	978	6	2.7	85	2	Q5VP42_ORYSA	Q5vp42 oryza sativ
906	6	2.7	66	2	Q6CH60_YARLI	Q6ch60 yarrowia 1i	979	6	2.7	85	2	Q5J7J4_ORYSA	Q5j7j4 oryza sativ
907	6	2.7	66	2	Q5CKL9_CRYHO	Q5ckl9 cryptospori	980	6	2.7	85	2	Q9WMG1_XANCV	Q9wmg1 xanthomonas

```

981 6 2.7 86 1 Y1279_IACAC O5f15 lactobacilli
982 6 2.7 86 2 Q4Y7K7_PLACH Q4Y7K7 plasmodium
983 6 2.7 86 2 Q648Y8_BACME Q648Y8 bacillus me
984 6 2.7 86 2 Q6M4H2_PAROW Q6M4H2 paracitand
985 6 2.7 87 2 Q4J1M0_AZCVI Q4J1M0 azotobacter
986 6 2.7 88 2 Q9T0Z7_9CAUD Q9T0Z7 bacterioph
987 6 2.7 88 2 Q6Q0Z8_HDV Q6Q0Z8 hepatitis d
988 6 2.7 88 2 Q6Q0Z9_HDV Q6Q0Z9 hepatitis d
989 6 2.7 88 2 Q6Q0R0_HDV Q6Q0R0 hepatitis d
990 6 2.7 88 2 Q74774_9HIV1 Q74774 human immun
991 6 2.7 89 2 Q6LY84_METMP Q6LY84 methanococ
992 6 2.7 89 2 Q7V5S9_PROMM Q7V5S9 prochloroc
993 6 2.7 89 2 Q35549_MUSSP Q35549 mus sprecus
994 6 2.7 90 2 Q7VUC9_BORPE Q7VUC9 bordelella
995 6 2.7 90 2 Q7W415_BORPA Q7W415 bordelella
996 6 2.7 90 2 Q6PDY8_MOUSE Q6PDY8 mus musculu
997 6 2.7 91 1 IIG1_CAEBL Q10060 caenorhadit
998 6 2.7 91 2 Q6S7Y5_9HIV1 Q6S7Y5 human immun
999 6 2.7 91 2 Q8US58_9HIV1 Q8US58 human immun
1000 6 2.7 92 2 Q5JW64_HUMAN Q5JW64 homo sapien

```

ALIGNMENTS

RESULT 1
Q922V0_MOUSE PRELIMINARY; PRT; 220 AA.

AC Q922V0;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Beta spectrin homolog.
GN Name=Spnb2; Synonyms=SLP1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

NCBI_TaxID=10090;

NUCLEOTIDE SEQUENCE.

RC STRAIN=C57 Black/6;

RA Cai T., Yu P., Mishra B., Mishra L.;

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

EMBL: AF016040; AAD01566.1; -; mRNA.

DR HSSP; Q15149; IMB8.

DR MGI; MGI:98388; Spnb2.

DR GO; GO:0005634; C:nucleus; IDA.

DR GO; GO:0005886; C:plasma membrane; IDA.

DR GO; GO:0005515; F:protein binding; IPI.

DR GO; GO:0007182; P:common-partner SMAD protein phosphorylation; IDA.

DR GO; GO:0007184; P:SMAD protein nuclear translocation; IDA.

DR InterPro; IPR001589; Actbind_actin.

DR InterPro; IPR001715; Calponin-like.

DR Pfam; PF00307; CH; 1.

DR SMART; SM00033; CH; 1.

DR PROSITE; PS00019; ACTININ_1; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

DR PROSITE; PS50021; CH; 1.

```

Db 121 GSHDIVGNHRLTTLLELVRRQOEERKRPPSPDPTKVSSEASQOMDTSKQDVS 180
QY 181 QNGIPAEQSPRVSRYRQYONYKNFNSRRTASHSNSGM 220
Db 181 QNGIPAEQSPRVSRYRQYONYKNFNSRRTASHSNSGM 220

```

RESULT 2

Q8C509_MOUSE PRELIMINARY; PRT; 155 AA.

AC Q8C509;

DT 01-MAR-2003 (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched

DE library, clone:A43008803 product:beta-spectrin 2, non-erythrocytic,

DE full insert sequence. (Fragment).

DE Name=Spnb2;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

NCBI_TaxID=10090;

NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Thymus;

RA MEDLINE=2108560; PubMed=11217851; DOI=10.1038/35055500;

RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Aikawa T., Hara A., Fukunishi Y., Kono H., Aachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,

RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Aeburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Mateuo Y., Nikaide I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombere P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoendach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,

RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlschut S.,

RA Hayashizaki Y.;

RA "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690 (2001).

NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Thymus;

RA The FANTOM Consortium;

RT "The RIKEN genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573 (2002).

NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Thymus;

RA MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,

RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to

RT prepare full-length cDNA libraries for rapid discovery of new genes.";

RL Genome Res. 10:1617-1630 (2000).

NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6J; TISSUE=Thymus;

RA MEDLINE=2108560; PubMed=11217851; DOI=10.1038/35055500;

RA Kawai T., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Aikawa T., Hara A., Fukunishi Y., Kono H., Aachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,

RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Thymus;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shihata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Katsunai T., Tashiro H., Itoh M.,
 RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachiiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer."; [6]
 RL Genome Res. 10:1757-1771(2000).
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Thymus;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kaikawa T.,
 RA Kato H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Nunazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Tova T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
 RL Submitted (Apr-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AK079842; BAC37762.1; -; mRNA.
 DR HSSP; Q15149; IMB8.
 DR MGI; MGI:98388; Spn2.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0005886; C:plasma membrane; IDA.
 DR GO; GO:0005513; F:protein binding; IPI.
 DR GO; GO:0007182; F:common-partner SMAD protein phosphorylation; IDA.
 DR GO; GO:0007184; F:SMAD protein nuclear translocation; IDA.
 DR InterPro: IPR001589; Actbind actin.
 DR InterPro: IPR001715; Calponin-like.
 DR Pfam: PF00307; CH; 1.
 DR SMART; SM00033; CH; 1.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 1.
 DR FT NON TER 155 155
 DR SQ SEQUENCE 155 AA; 18020 MW; E9F58F25CA6BDF8 CRC64;
 Query Match 60.5%; Score 133; DB 2; Length 155;
 Best Local Similarity 100.0%; Pred. No. 1.3e-128;
 Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELQRTSSISGPLSPAYTGVPPYNNQLEGRFKQLQDEREAVOKTFTKWNSHLARVSC 60
 DB 1 MELQRTSSISGPLSPAYTGVPPYNNQLEGRFKQLQDEREAVOKTFTKWNSHLARVSC 60
 QY 61 RITDLYTDLRGRMLIKLEVLSEGRLPKPTKGRMRHCLENVKALQFLKEQVRHLENN 120
 DB 61 RITDLYTDLRGRMLIKLEVLSEGRLPKPTKGRMRHCLENVKALQFLKEQVRHLENN 120
 QY 121 GSHDIVDGNHRLT 133
 DB 121 GSHDIVDGNHRLT 133
 RESULT 3
 ID Q5SOL9 MOUSE PRELIMINARY; PRT; 2154 AA.
 AC Q5SOL9;
 DT 01-FEB-2005 (TRENBLrel. 29, Created)
 DT 01-FEB-2005 (TRENBLrel. 29, Last sequence update)
 DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
 DE Spectrin beta 2.
 GN Name=Spn2; ORFNames=RP23-189P1.1-002;
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN
 RP NUCLEOTIDE SEQUENCE.
 RA Phillimore B.;
 RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Troman A.;
 RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AL731792; CAI24366.1; -; Genomic DNA.
 DR EMBL; AL672225; CAI25429.1; -; Genomic DNA.
 DR EMBL; AL731792; CAI25429.1; JOINED; Genomic DNA.
 DR Ensembl; ENSMUSG0000020315; Mus musculus.
 DR GO; GO:0003779; F:actin binding; IEA.
 DR InterPro: IPR001589; Actin actin bd.
 DR InterPro: IPR001715; Calponin act_bd.
 DR InterPro: IPR002017; Spectrin.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00435; Spectrin; 17.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00150; SPEC; 17.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR SQ SEQUENCE 2154 AA; 251156 MW; 5128A840C3FD28AC CRC64;
 Query Match 60.5%; Score 133; DB 2; Length 2154;
 Best Local Similarity 100.0%; Pred. No. 1.6e-127;
 Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELQRTSSISGPLSPAYTGVPPYNNQLEGRFKQLQDEREAVOKTFTKWNSHLARVSC 60
 DB 1 MELQRTSSISGPLSPAYTGVPPYNNQLEGRFKQLQDEREAVOKTFTKWNSHLARVSC 60
 QY 61 RITDLYTDLRGRMLIKLEVLSEGRLPKPTKGRMRHCLENVKALQFLKEQVRHLENN 120
 DB 61 RITDLYTDLRGRMLIKLEVLSEGRLPKPTKGRMRHCLENVKALQFLKEQVRHLENN 120
 QY 121 GSHDIVDGNHRLT 133
 DB 121 GSHDIVDGNHRLT 133
 RESULT 4
 ID O8IX99 HUMAN PRELIMINARY; PRT; 2155 AA.
 AC O8IX99;
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Beta-spectrin 2 isoform 2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=21521960; PubMed=1165863; DOI=10.1385/JMN:17.1:59;
 RA Chen Y., Yu P., Lu D., Tagle D.A., Cai T.;
 RT "A novel isoform of beta-spectrin II localizes to cerebellar Purkinje-
 RT cell bodies and interacts with neurofibromatosis type 2 gene product
 RT schwannomin.";
 RL J. Mol. Neurosci. 17:59-70(2001).
 DR EMBL; AF37441; AA015362.1; -; mRNA.
 DR HSSP; Q01082; IBKR.
 DR Ensembl; ENSG00000115306; Homo sapiens.
 DR GO; GO:0005856; C:cytoskeleton; IEA.
 DR GO; GO:0003779; F:actin binding; IEA.

DR InterPro; IPR001589; Acbind_actuin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00335; Spectrin; 17.
DR SMART; SM00033; CH; 2.
DR SMART; SM00150; SPEC; 17.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS00021; CH; 2.
SEQUENCE 2155 AA; 251418 MW; 20A91A28AC0FBED CRC64;

Query Match 60.5%; Score 133; DB 2; Length 2155;
Best Local Similarity 100.0%; Pred. No. 1.6e-127;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELQRTSSISGPLSPAYTQGVYNNQLGGRFKQLODEREAVQKTFKVNSHLARVSC 60
DB 1 MELQRTSSISGPLSPAYTQGVYNNQLGGRFKQLODEREAVQKTFKVNSHLARVSC 60
QY 61 RTTDTYTDRDGRMLIKLEVLGSRGLPKPTKGRMRHCLENVDKALQFLKEQVHLENN 120
DB 61 RTTDTYTDRDGRMLIKLEVLGSRGLPKPTKGRMRHCLENVDKALQFLKEQVHLENN 120
QY 121 GSHDIVDGNHRLT 133
DB 121 GSHDIVDGNHRLT 133

RESULT 5
Q9QWJ7 MOUSE PRELIMINARY; PRT; 2154 AA.
ID Q9QWJ7
AC Q9QWJ7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Non-erythrocyte beta spectrin.
GN Name=Spnb2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CS7 Black/6;
RA Cai T., Yu P., Mishra B., Monga P.S.P., Mishra L.;
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.
DR HSSP; Q01082; 1BKR.
DR MGI; MGI:968389; Spnb2.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0005515; F:protein binding; IPT.
DR GO; GO:0007184; P:common-partner SMAD protein phosphorylation; IDA.
DR GO; GO:0007184; P:SMAD protein nuclear translocation; IDA.
DR InterPro; IPR001589; Acbind_actuin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00335; Spectrin; 17.
DR SMART; SM00033; CH; 2.
DR SMART; SM00150; SPEC; 17.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS00021; CH; 2.
SEQUENCE 2154 AA; 250928 MW; 872CFEC9C332152C CRC64;

Query Match 56.4%; Score 124; DB 2; Length 2154;
Best Local Similarity 100.0%; Pred. No. 3.2e-118;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 SGPLSPAYTQGVYNNQLGGRFKQLODEREAVQKTFKVNSHLARVSCRTTDTYTDL 69

DB 10 SGPLSPAYTQGVYNNQLGGRFKQLODEREAVQKTFKVNSHLARVSCRTTDTYTDL 69
QY 70 RDGRMLIKLEVLGSRGLPKPTKGRMRHCLENVDKALQFLKEQVHLENNGSHDIVDGN 129
DB 70 RDGRMLIKLEVLGSRGLPKPTKGRMRHCLENVDKALQFLKEQVHLENNGSHDIVDGN 129
QY 130 HRLT 133
DB 130 HRLT 133

RESULT 6
Q6V0G7 CANFA PRELIMINARY; PRT; 149 AA.
ID Q6V0G7
AC Q6V0G7
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Beta-spectrin (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC Canis.
OX NCBI_TaxID=9615;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mishra S., Sabbah H.N., Undrovinae A.I.;
RL Submitted (Jul-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY356552; AA063200.1; mRNA.
DR Ensembl; ENSGAFG00000002744; Canis familiaris.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR001589; Acbind_actuin.
DR InterPro; IPR001715; Calponin-like.
DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00021; CH; 1.
FT NON TER 149
FT 149
SEQUENCE 149 AA; 17412 MW; 2281545132DD85CC CRC64;

Query Match 44.1%; Score 97; DB 2; Length 149;
Best Local Similarity 100.0%; Pred. No. 1.9e-91;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 DERBAVOKTFTKVNSHLARVSCRTTDTYTDLRGRMLIKLEVLGSRGLPKPTKGRMR 96
DB 50 DERBAVOKTFTKVNSHLARVSCRTTDTYTDLRGRMLIKLEVLGSRGLPKPTKGRMR 109
QY 97 IHCLENVDKALQFLKEQVHLENNGSHDIVDGNHRLT 133
DB 110 IHCLENVDKALQFLKEQVHLENNGSHDIVDGNHRLT 146

RESULT 7
Q5DTR4 MOUSE PRELIMINARY; PRT; 1290 AA.
ID Q5DTR4
AC Q5DTR4
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE MKIAA4049 protein (Fragment).
GN Name=Spnb2; Synonyms=mkIAA4049;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Nagase T., Ohara O., Koga H.;

RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene.
 RT The Complete Nucleotide Sequences of Mouse KIAA-homologues cDNAs
 RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
 RT Sampled from Size-Fractionated Libraries.";
 RL Submitted (FEB-2005) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AK220455; BAD90485.1; -; mRNA.
 DR MGI; MGI:98388; Snpb2.
 DR GO; GO:0005634; C:nucleus; IDA.
 DR GO; GO:0005886; C:plasma membrane; IDA.
 DR GO; GO:0005515; P:protein binding; IPI.
 DR GO; GO:0007184; P:common-partner SMAD protein phosphorylation; IDA.
 DR GO; GO:0007184; P:SMAD protein nuclear translocation; IDA.
 DR InterPro; IPR001589; Acbind actin.
 DR InterPro; IPR001715; Calponin-like.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00435; Spectrin; 9.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00150; SPEC; 9.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 FT NON_TER 1 1
 SQ SEQUENCE 1290 AA; 150490 MW; D9E1CD33E6AB299B CRC64;
 Query Match 44.1%; Score 97; DB 2; Length 1290;
 Best Local Similarity 100.0%; Pred. No. 1.5e-90;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 37 DEREAVQKTKTKWNSHARVSCRITDLYDLRDGRMLIKLEVLGSGRLPKPTKGMNR 96
 DB 59 DEREAVQKTKTKWNSHARVSCRITDLYDLRDGRMLIKLEVLGSGRLPKPTKGMNR 118
 QY 97 IHCLENVDKALQFLKEQVHLENMGSHDIVDGNHRLT 133
 DB 119 IHCLENVDKALQFLKEQVHLENMGSHDIVDGNHRLT 155
 RESULT 8
 Q6XD99_RAT PRELIMINARY; PRT; 2358 AA.
 AC Q6XD99;
 DT 05-JUL-2004 (Tremblrel. 27, Created)
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE Non-erythroid spectrin beta.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Rattus;
 OC NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Chen Y., Huang C.-H.;
 RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 PH domain.
 DR EMBL; AY238344; AA002380.1; -; mRNA.
 DR HSSP; Q01082; 1A2.
 DR SMR; Q6XD99; 2194-2299.
 DR Ensembl; ENSRMG00000005434; Rattus norvegicus.
 DR GO; GO:0003779; F:actin binding; IEA.
 DR InterPro; IPR001589; Actnin actin bd.
 DR InterPro; IPR001715; Calponin act_bd.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR001605; Spectrin_PH.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00435; Spectrin; 9.
 DR Pfam; PF00633; CH; 2.
 DR PRINTS; SM00033; SPECTRINPH.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00150; SPEC; 17.

DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS50003; PH DOMAIN; 1.
 KW Actin capring; Actin-binding; Cytoskeleton.
 SQ SEQUENCE 2358 AA; 273586 MW; 6252854560453C8 CRC64;
 Query Match 44.1%; Score 97; DB 2; Length 2358;
 Best Local Similarity 100.0%; Pred. No. 2.7e-90;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 37 DEREAVQKTKTKWNSHARVSCRITDLYDLRDGRMLIKLEVLGSGRLPKPTKGMNR 96
 DB 50 DEREAVQKTKTKWNSHARVSCRITDLYDLRDGRMLIKLEVLGSGRLPKPTKGMNR 109
 QY 97 IHCLENVDKALQFLKEQVHLENMGSHDIVDGNHRLT 133
 DB 110 IHCLENVDKALQFLKEQVHLENMGSHDIVDGNHRLT 146
 RESULT 9
 SPTB2_MOUSE STANDARD; PRT; 2363 AA.
 AC Q62261;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Spectrin beta chain, brain I (Spectrin, non-erythroid beta chain I)
 DE (Beta-II spectrin) (Fodrin beta chain)
 GN Name=Spbchl; Synonyms=Spnb-2, Spnb2, Spb2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BALB/c;
 RA Ma Y., Zimmer W.B., Riederer B.M., Goodman S.R.;
 RT "The complete amino acid sequence for brain beta spectrin (beta fodrin): relationship to globin sequences.";
 RL Brain Res. Mol. Brain Res. 18:87-89(1993).
 RN [2]
 RP PHOSPHORYLATION SITES SER-1918, SER-2115, SER-2137 AND SER-2168.
 RA PubMed=15572359; DOI=10.1074/jbc.M411220200;
 RA Collins M.O., Yu L., Caba M.P., Husi H., Campuzano I.,
 RA Blackstock W.P., Choudhary J.S., Grant S.G.;
 RT "Proteomic analysis of in vivo phosphorylated synaptic proteins.";
 RL J. Biol. Chem. 280:5972-5982(2005).
 RN [3]
 RP STRUCTURE BY NMR OF 2199-2304.
 RA MEDLINE=94268558; PubMed=8208297; DOI=10.1038/36967580;
 RA Macias M.J., Musacchio A., Ponstingl H., Nilges M., Saraste M.,
 RA Osochkinat H.;
 RL "Structure of the pleckstrin homology domain from beta-spectrin.";
 RL Nature 369:675-677(1994).
 RN [4]
 RP STRUCTURE BY NMR OF 2199-2304.
 RA MEDLINE=97342712; PubMed=1999409; DOI=10.1006/jmbi.1997.1044;
 RA Nilges M., Macias M.J., O'Donoghue S.I., Osochkinat H.;
 RT "Automated NOESY interpretation with ambiguous distance restraints:
 the refined NMR solution structure of the pleckstrin homology domain
 from beta-spectrin.";
 RL J. Mol. Biol. 269:408-422(1997).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 2199-2304.
 RA MEDLINE=96030773; PubMed=7588597;
 RA Hyvoenen M., Macias M.J., Nilges M., Osochkinat H., Saraste M.,
 RA Wilmanns M.;
 RT "Structure of the binding site for inositol phosphates in a PH
 domain.";
 RL EMBO J. 14:4676-4681(1995).

CC -1- FUNCTION: Podrin, which seems to be involved in secretion,
 CC interacts with calmodulin in a calcium-dependent manner and is
 CC this candidate for the calcium-dependent movement of the
 CC cytoskeleton at the membrane.
 CC -1- SUBUNIT: Like erythrocyte spectrin, the spectrin-like proteins are
 CC capable to form dimers which can further associate to tetramers.
 CC -1- SIMILARITY: Belongs to the spectrin family.
 CC -1- SIMILARITY: Contains 2 CH (calponin-homology) domains.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL, M74773; AAC2040.1; -; mRNA.
 CC PDB, 1BPN; X-ray; @=2199-2304.
 CC DR Intact; O62261; -;
 CC DR Ensembl; ENSMUSG0000020315; Mus musculus.
 CC MGI; MGI:98388; Spn2.
 CC DR GO; GO:0005634; C:nucleus; IDA.
 CC DR GO; GO:0005886; C:plasma membrane; IDA.
 CC DR GO; GO:0005815; F:protein binding; IPI.
 CC DR GO; GO:0007182; P:common-partner SMAD protein phosphorylation; IDA.
 CC DR GO; GO:0007184; P:SMAD protein nuclear translocation; IDA.
 CC DR InterPro; IPR001589; Actinin actin bd.
 CC DR InterPro; IPR001715; Calponin_act_bd.
 CC DR InterPro; IPR011849; PH.
 CC DR InterPro; IPR02017; Spectrin.
 CC DR InterPro; IPR02017; Spectrin.
 CC DR Pfam; PF00307; CH_2.
 CC DR Pfam; PF00169; PH; 1.
 CC DR Pfam; PF00435; Spectrin; 16.
 CC DR PRINTS; PR00683; SPECTRINPH.
 CC DR PROSITE; PS00019; ACTININ_1; 1.
 CC DR PROSITE; PS00020; ACTININ_2; 1.
 CC DR PROSITE; PS50021; CH; 2.
 CC DR PROSITE; PS50003; PH_DOMAIN; 1.
 CC DR 3D-structure; Actin capping; Actin-binding; Calmodulin-binding;
 CC Cytoskeleton; Membrane; Phosphorylation; Repeat.
 CC KM DOMAIN 1 275 Actin-binding (By similarity).
 CC FT DOMAIN 54 158 CH 1.
 CC FT REPEAT 173 275 Spectrin 1.
 CC FT REPEAT 276 384 Spectrin 2.
 CC FT REPEAT 385 498 Spectrin 3.
 CC FT REPEAT 499 608 Spectrin 4.
 CC FT REPEAT 609 714 Spectrin 5.
 CC FT REPEAT 715 819 Spectrin 6.
 CC FT REPEAT 820 925 Spectrin 7.
 CC FT REPEAT 926 1032 Spectrin 8.
 CC FT REPEAT 1033 1140 Spectrin 9.
 CC FT REPEAT 1140 1245 Spectrin 10.
 CC FT REPEAT 1246 1350 Spectrin 11.
 CC FT REPEAT 1351 1462 Spectrin 12.
 CC FT REPEAT 1463 1562 Spectrin 13.
 CC FT REPEAT 1563 1668 Spectrin 14.
 CC FT REPEAT 1669 1775 Spectrin 15.
 CC FT REPEAT 1776 1881 Spectrin 16.
 CC FT REPEAT 1882 1987 Spectrin 17.
 CC FT REPEAT 1988 2132 Spectrin 17.
 CC FT DOMAIN 2136 2306 PH.
 CC FT MOD_RES 2196 2306 Phosphoserine.
 CC FT MOD_RES 2198 1918 Phosphoserine.
 CC FT MOD_RES 2115 2115 Phosphoserine (By similarity).
 CC FT MOD_RES 2127 2127 Phosphoserine (By similarity).
 CC FT MOD_RES 2137 2137 Phosphoserine.
 CC FT MOD_RES 2164 2164 Phosphoserine (By similarity).
 CC FT MOD_RES 2168 2168 Phosphoserine.
 CC FT STRAND 2200 2209
 CC FT TURN 2212 2213

FT STRAND 2215 2215
 FT STRAND 2222 2229
 FT TURN 2230 2231
 FT STRAND 2232 2236
 FT HELIX 2239 2244
 FT TURN 2245 2245
 FT STRAND 2254 2255
 FT TURN 2257 2258
 FT STRAND 2260 2263
 FT STRAND 2273 2277
 FT TURN 2279 2280
 FT STRAND 2283 2287
 FT HELIX 2291 2303
 SQ SEQUENCE 2363 AA; 274422 MW; 64C9ABD26B8C7B8 CRC64;
 Query Match 44.1%; Score 97; DB 1; Length 2363;
 Best Local Similarity 100.0%; Pred. No. 2,7e-90;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 37 DERAQVQKPTTKVNSHARVSCRTDLYTDLRDGRMLIKLLEVLGGERLPKPTKGRMR 96
 DB 50 DERAQVQKPTTKVNSHARVSCRTDLYTDLRDGRMLIKLLEVLGGERLPKPTKGRMR 109
 QY 97 IHCLENVDKALQFLKEQVHLNNGSHDIYDGNRLT 133
 DB 110 IHCLENVDKALQFLKEQVHLNNGSHDIYDGNRLT 146
 RESULT 10
 Q55QJ8_MOUSE Q55QJ8_MOUSE PRELIMINARY; PRT; 2363 AA.
 AC Q55QJ8;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Spectrin beta 2.
 GN Name=Spn2; ORFNames=RP23-189P1.1-001;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OC NCBI_taxid=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Philimore B.;
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Tyomans A.;
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC EMBL; AL731792; CAI24367.1; -; Genomic DNA.
 CC EMBL; AL672225; CAI24367.1; JOINED; Genomic DNA.
 CC EMBL; AL731792; CAI25430.1; JOINED; Genomic DNA.
 CC SMR; Q55QJ8; 2199-2304.
 CC DR Ensembl; ENSMUSG0000020315; Mus musculus.
 CC DR GO; GO:0003779; F:actin binding; IEA.
 CC DR InterPro; IPR001589; Actinin actin bd.
 CC DR InterPro; IPR001715; Calponin_act_bd.
 CC DR InterPro; IPR011849; PH.
 CC DR InterPro; IPR011993; PH type.
 CC DR InterPro; IPR02017; Spectrin.
 CC DR InterPro; IPR01605; Spectrin_PH.
 CC DR Pfam; PF00307; CH; 2.
 CC DR Pfam; PF00169; PH; 1.
 CC DR Pfam; PF00435; Spectrin; 17.
 CC DR PRINTS; PR00683; SPECTRINPH.
 CC DR SMART; SM00033; CH; 2.
 CC DR SMART; SM00233; PH; 1.
 CC DR SMART; SM00150; SPEC; 17.
 CC DR PROSITE; PS00019; ACTININ_1; 1.
 CC DR PROSITE; PS00020; ACTININ_2; 1.

DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS50003; PH DOMAIN; 1.
 KW Actin capping; Actin-binding; Cytoskeleton.
 SQ SEQUENCE 2363 AA; 274223 MW; 2213620545E4B8C CRC64;

Query Match: 44.1%; Score 97; DB 2; Length 2363;
 Best Local Similarity 100.0%; Pred. No. 2,7e-90;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 DERENVOKTTTQWNSHARVSCRTITLTYDLRGRMLIKLEVLSEGRLPKPKGMR 96
 DB 50 DERENVOKTTTQWNSHARVSCRTITLTYDLRGRMLIKLEVLSEGRLPKPKGMR 109
 QY 97 IHCLENVDKALQFLKEQVHLENMGSHDIVDGNHRLT 133
 DB 110 IHCLENVDKALQFLKEQVHLENMGSHDIVDGNHRLT 146

RESULT 11
 SPTB2 HUMAN STANDARD; PRT; 2364 AA.
 AC Q01082; Q06083; Q16057;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Spectrin beta chain, brain 1 (Spectrin, non-erythroid beta chain 1)
 DE (Beta-II spectrin) (Fodrin beta chain)
 GN Name=SPTB1; Synonyms=SPTB2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 NCBI_TaxID=9606;
 RX NUCLEOTIDE SEQUENCE (ISOFORM LONG).
 RC TISSUE=Brain;
 RX MEDLINE=92406787; PubMed=1527002;
 RA Hu R.J., Watanabe M., Bennett V.;
 RT "Characterization of human brain cDNA encoding the general isoform of
 RT beta-spectrin.";
 RL J. Biol. Chem. 267:18715-18722(1992).
 RN [2]
 RP NUCLEOTIDE SEQUENCE OF 293-1544.
 RX MEDLINE=94010920; PubMed=8406479;
 RA Chang J.G., Scarpa A., Eddy R.L., Byers M.G., Harris A.S.,
 RA Morrow J.S., Watkins P., Shows T.B., Forget B.G.;
 RT "Cloning of a portion of the chromosomal gene and cDNA for human beta-
 RT fodrin, the nonerythroid form of beta-spectrin.";
 RL Genomics 17:287-293(1993).
 RN [3]
 RP NUCLEOTIDE SEQUENCE OF 2087-2168 (ISOFORM SHORT).
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=20267884; PubMed=10806113;
 RA Hayes N.V.L., Scott C., Heerkes E., Oshanian V., Maggs A.M.,
 RA Pinder J.C., Kordeli E., Baines A.J.;
 RT "Identification of a novel C-terminal variant of beta1 spectrin: two
 RT isoforms of beta1 spectrin have distinct intracellular locations and
 RT activities.";
 RL J. Cell Sci. 113:2023-2034(2000).
 RN [4]
 RP PHOSPHORYLATION SITES SRR-2128; SRR-2138; SRR-2165 AND SRR-2169.
 RX PubMed15502935; DOI=10.1073/pnas.0404720101;
 RA Beausoleil S.A., Jedrychowski M., Schwartz D., Elias J.E., Villen J.,
 RA Li J., Cohn M.A., Cantley L.C., Gygi S.P.;
 RT "Large-scale characterization of HeLa cell nuclear phosphoproteins.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:12130-12135(2004).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 173-280.
 RX MEDLINE=97307247; PubMed=9164454;
 RA Carugo K.D., Bannet S., Saraste M.;
 RT "Crystal structure of a calponin homology domain.";
 RL Nat. Struct. Biol. 4:175-179(1997).
 RN [6]

RP X-RAY CRYSTALLOGRAPHY (1.1 ANGSTROMS) OF 173-281.
 RX MEDLINE=99036861; PubMed=9817844; DOI=10.1016/S0969-2126(98)00141-5;
 RA Bannet S., Saraste M., Carugo K.D.;
 RT "Structural comparisons of calponin homology domains: implications for
 RT actin binding.";
 RL Structure 6:1419-1431(1998).
 CC -1- FUNCTION: Fodrin, which seems to be involved in secretion,
 CC interacts with calmodulin in a calcium-dependent manner and is
 CC thus candidate for the calcium-dependent movement of the
 CC cytoskeleton at the membrane.
 CC -1- SUBUNIT: Like erythrocyte spectrin, the spectrin-like proteins are
 CC capable to form dimers which can further associate to tetramers.
 CC The short form cannot bind to the axonal protein fodaxin.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=Q01082-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=Q01082-2; Sequence=VSP_000720, VSP_000721;
 CC -1- SIMILARITY: Belongs to the spectrin family.
 CC -1- SIMILARITY: Contains 2 CH (calponin-homology) domains.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -1- SIMILARITY: Contains 17 spectrin repeats.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; M96803; AAA60580.1; -; mRNA.
 CC EMBL; S65762; AAB28324.1; -; mRNA.
 CC EMBL; AJ005634; CAA06678.1; -; mRNA.
 CC EMBL; AJ238723; CAB91088.1; -; Genomic_DNA.
 CC PIR; A44159; A44159.
 CC PIR; A47213; A47213.
 CC PDB; 1AA2; X-ray; @=173-280.
 CC PDB; 1BKR; X-ray; A=172-280.
 CC SMR; Q01082; 2200-2305.
 CC InAcct; Q01082; -;
 CC EMBL; ENSG00000115306; Homo sapiens.
 CC HGNC; HGNC:11275; SPTB1.
 CC MIM; 182790; -;
 CC GO; GO:0008091; C:spectrin; TAS.
 CC GO; GO:0003779; F:actin binding; TAS.
 CC GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
 CC InterPro; IPR001589; Actinin actin bd.
 CC InterPro; IPR001715; Calponin_act_bd.
 CC InterPro; IPR001849; PH.
 CC InterPro; IPR011993; PH_type.
 CC InterPro; IPR002017; Spectrin.
 CC InterPro; IPR001605; Spectrin_PH.
 CC Pfam; PF00307; PH; 2.
 CC Pfam; PF00169; PH; 1.
 CC Pfam; PF00435; SpectrinPH.
 CC PRINTS; PR00683; SPECTRINPH.
 CC PROSITE; PS00019; ACTININ_1; 1.
 CC PROSITE; PS00020; ACTININ_2; 1.
 CC PROSITE; PS50021; CH; 2.
 CC PROSITE; PS50003; PH DOMAIN; 1.
 CC 3D-structure; Actin capping; Actin-binding; Alternative splicing;
 CC Calmodulin-binding; Cytoskeleton; Membrane; Phosphorylation; Repeat.
 CC Domain-binding (By similarity).
 CC DOMAIN 1 275
 CC CH 1.
 CC DOMAIN 2 158
 CC CH 2.
 CC DOMAIN 3 173
 CC CH 3.
 CC REPEAT 1 276
 CC Spectrin 1.
 CC REPEAT 2 385
 CC Spectrin 2.
 CC REPEAT 3 499
 CC Spectrin 3.
 CC REPEAT 4 609
 CC Spectrin 4.
 CC REPEAT 5 715
 CC Spectrin 5.
 CC REPEAT 6 820
 CC Spectrin 6.
 CC REPEAT 7 926
 CC Spectrin 7.
 CC REPEAT 8 1033
 CC Spectrin 8.

```

FT REPEAT 1140 1245 Spectrin 9.
FT REPEAT 1246 1350 Spectrin 10.
FT REPEAT 1351 1462 Spectrin 11.
FT REPEAT 1463 1562 Spectrin 12.
FT REPEAT 1563 1668 Spectrin 13.
FT REPEAT 1669 1775 Spectrin 14.
FT REPEAT 1776 1881 Spectrin 15.
FT REPEAT 1882 1987 Spectrin 16.
FT REPEAT 1988 2133 Spectrin 17.
FT DOMAIN 2197 2307 Phosphoserine (By similarity).
FT MOD_RES 2198 1918 Phosphoserine (By similarity).
FT MOD_RES 2198 2115 Phosphoserine (By similarity).
FT MOD_RES 2198 2128 Phosphoserine.
FT MOD_RES 2138 2138 Phosphoserine.
FT MOD_RES 2165 2165 Phosphoserine.
FT MOD_RES 2169 2169 Phosphoserine.
FT VARSPLIC 2141 2168 MAEYDTSBMVNGATEORTSSKESPP -> VSTRSOTQY
NTKNTSKRTASDQPSGL (1in Isoform Short).
FT VARSPLIC 2169 2364 Missing (in Isoform Short).
FT VARSPLIC 2169 2364 /FTid=VSP_000720.
FT CONFLICT 1411 1411 H -> D (1in Ref. 2).
FT HELIX 174 186 /FTid=VSP_000721.
FT TURN 187 189
FT TURN 191 192
FT TURN 200 202
FT TURN 203 204
FT HELIX 206 215
FT HELIX 217 219
FT HELIX 222 224
FT TURN 227 228
FT HELIX 230 245
FT HELIX 253 256
FT HELIX 263 277
FT TURN 278 278
SQ SEQUENCE 2364 AA; 274631 MW; 0122DB3DF25872CC CRC64;

Query Match 44.1%; Score 97; DB 1; Length 2364;
Best Local Similarity 100.0%; Pred. No. 2,7e-90;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 DERBAVQKTFKTVNNSHLARVSCRTIDLYTLDRGRLIKLLEVLSGERLPKPTKGRMR 96
DB 50 DERBAVQKTFKTVNNSHLARVSCRTIDLYTLDRGRLIKLLEVLSGERLPKPTKGRMR 109
QY 97 IHCLENVDKALQFLKEQRVHLENMGSHDIVDGNHRLT 133
DB 110 IHCLENVDKALQFLKEQRVHLENMGSHDIVDGNHRLT 146

RESULT 12
OS9ER3_HUMAN PRELIMINARY; PRT; 2377 AA.
ID OS9ER3_HUMAN
AC OS9ER3;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE Spectrin, beta, non-erythrocytic 1 isoform 1 variant (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
OX NCB1
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Brain;
RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AB209748; BAD92985.1; -; mRNA.
DR SMR; OS9ER3; 2213-2318.

```

```

DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR001589; Actin actin_bd.
DR InterPro; IPR001715; Calponin_act_bd.
DR InterPro; IPR001849; PH.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001605; Spectrin_PH.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00435; Spectrin_17.
DR PRINTS; PR00683; SPECTRINPH.
DR SMART; SM00033; CH; 2.
DR SMART; SM00150; SPEC; 17.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
DR PROSITE; PS50003; PH DOMAIN; 1.
DR Actin capding; Actin-binding.
FT NON_TER 1
SQ SEQUENCE 2377 AA; 276170 MW; E497320876298877 CRC64;

Query Match 44.1%; Score 97; DB 2; Length 2377;
Best Local Similarity 100.0%; Pred. No. 2,8e-90;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 DERBAVQKTFKTVNNSHLARVSCRTIDLYTLDRGRLIKLLEVLSGERLPKPTKGRMR 96
DB 63 DERBAVQKTFKTVNNSHLARVSCRTIDLYTLDRGRLIKLLEVLSGERLPKPTKGRMR 122
QY 97 IHCLENVDKALQFLKEQRVHLENMGSHDIVDGNHRLT 133
DB 123 IHCLENVDKALQFLKEQRVHLENMGSHDIVDGNHRLT 159

RESULT 13
OS3R89_HUMAN PRELIMINARY; PRT; 2314 AA.
ID OS3R89_HUMAN
AC OS3R89;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Spectrin, beta, non-erythrocytic 1 isoform 1 variant (Fragment).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
OX NCB1
RN NUCLEOTIDE SEQUENCE.
RP KOZLOVICZ A., SPALDING L., MULVANEY E.;
RA "The sequence of Homo sapiens BAC clone RP11-564H16."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA WATERSTON R.H.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA WATERSTON R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RN NUCLEOTIDE SEQUENCE.
RA WILSON R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AC093110; AA124229.1; -; Genomic_DNA.
DR SMR; OS3R89; 2150-2255.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR001589; Actin actin_bd.
DR InterPro; IPR001715; Calponin_act_bd.
DR InterPro; IPR001849; PH.
DR InterPro; IPR002017; Spectrin.

```

DR InterPro: IPR001605; Spectrin_PH.
DR Pfam: PF00307; CH; 2.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00435; Spectrin; 17.
DR PRINTS: PR00683; SpectrinPH.
DR SMART: SM00033; CH; 2.
DR SMART: SM00233; PH; 1.
DR SMART: SM00150; SPEC; 17.
DR PROSITE: PS00019; ACTININ_1; 1.
DR PROSITE: PS00020; ACTININ_2; 1.
DR PROSITE: PS50021; CH; 2.
DR PROSITE: PS50003; PH DOMAIN; 1.
DR Actin capping; Actin-binding; Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 2314 AA; 268716 MW; 9C49E79CFAD22081 CRC64;

Query Match 43.6%; Score 96; DB 2; Length 2314;
Best Local Similarity 100.0%; Pred. No. 2,9e-89;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 EREAVQKTFKRNKNSHARVSCRTDLYTDLRQGRMLIKLEVLSEGRLEPKPTGRRRI 97
Db 1 EREAVQKTFKRNKNSHARVSCRTDLYTDLRQGRMLIKLEVLSEGRLEPKPTGRRRI 60

Qy 98 HCLENVKALQFLKEQVHLENMGSHDIVDGNHRLT 133
Db 61 HCLENVKALQFLKEQVHLENMGSHDIVDGNHRLT 96

RESULT 14
ID Q502M3 BRARE PRELIMINARY; PRT; 977 AA.
AC Q502M3
DT 13-SEP-2005 (TrEMBLrel. 31; Created)
DT 13-SEP-2005 (TrEMBLrel. 31; Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31; Last annotation update)
DE LOC553451 protein (Fragment).
GN Breahevdanio reio (zebrafish) (Danio rerio).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
NCBI_TaxID=7955;
RX NUCLEOTIDE SEQUENCE.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., USCIN T.B., Toshitoki S., Carinini P., Prange C.,
RA Baha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosa S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huijck S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting R.W., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzyzinski M.T., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Olfactory epithelium;
RG NIH MGC Project;
RL Submitted (MAY-2005) to the EMBL/Genbank/DBJ databases.
EMBL: BC095643; AA095643.1; -, mRNA.

FT NON TER 977 977
SQ SEQUENCE 977 AA; 113881 MW; 5CC81C094A58BD01 CRC64;

Query Match 30.0%; Score 66; DB 2; Length 977;
Best Local Similarity 100.0%; Pred. No. 1,2e-58;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 DLRDGRMLIKLEVLSEGRLEPKPTGRRRIHCLNVKALQFLKEQVHLENMGSHDIVD 127
Db 80 DLRDGRMLIKLEVLSEGRLEPKPTGRRRIHCLNVKALQFLKEQVHLENMGSHDIVD 139

Qy 128 GNRHL 132
Db 140 GNRHL 145

RESULT 15
ID Q4REP5 TETNG PRELIMINARY; PRT; 2235 AA.
AC Q4REP5
DT 13-SEP-2005 (TrEMBLrel. 31; Created)
DT 13-SEP-2005 (TrEMBLrel. 31; Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31; Last annotation update)
DE Chromosome 10 SCAP15123, whole genome shotgun sequence.
GN ORFNames=STENG00035677001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acantthomorphia; Acantthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
NCBI_TaxID=99883;
RX NUCLEOTIDE SEQUENCE.
RC Genoscope; Whitehead Institute Centre for Genome Research;
RG Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: CA001015123; CA013137.1; -, Genomic DNA.
SQ SEQUENCE 2235 AA; 258437 MW; 04A0C7751E04AED9 CRC64;

Query Match 29.5%; Score 65; DB 2; Length 2235;
Best Local Similarity 100.0%; Pred. No. 2,9e-57;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 DLRDGRMLIKLEVLSEGRLEPKPTGRRRIHCLNVKALQFLKEQVHLENMGSHDIVD 127
Db 81 DLRDGRMLIKLEVLSEGRLEPKPTGRRRIHCLNVKALQFLKEQVHLENMGSHDIVD 140

Qy 128 GNRHL 132
Db 141 GNRHL 145

RESULT 16
Q90W73_gmuri

ID O9QW73 9MURI PRELIMINARY; PRT; 332 AA.
 AC O9QW73;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Beta-spectrin subunit (Fragment).
 OS Mus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 NC NCB1_TaxID=10095;
 RX NUCLEOTIDE SEQUENCE.
 RP MEDLINE=92076240; PubMed=1742606; DOI=10.1016/0361-9230(91)90066-S;
 RA Zimmer W.E., Ma Y.P., Goodman S.R.;
 RT "Identification of a mouse brain beta-spectrin cDNA and distribution
 of its mRNA in adult tissues."
 RL Brain Res. Bull. 27:187-193(1991).
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF00435; Spectrin; 3.
 DR SMART; SM00150; SPEC; 2.
 FT NON TER 1
 FT SEQUENCE 332 AA; 38192 MW; 5C9F36C840DDEEC3 CRC64;
 SQ
 Query Match 28.2%; Score 62; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 5,8e-55;
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 131 RITTELEVRROQEEERKRRPPDPPTKVSSEASQMDTSQGDVSONGLPAEAGS 190
 DB 204 RITTELEVRROQEEERKRRPPDPPTKVSSEASQMDTSQGDVSONGLPAEAGS 263
 QY 191 PR 192
 DB 264 PR 265
 RESULT 17
 O8B035_MOUSE PRELIMINARY; PRT; 497 AA.
 ID O8B035;
 AC O8B035;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
 enriched library, clone:DI30061A22 product:beta-spectrin 2, non-
 erythrocytic, full insert sequence. (Fragment).
 GN Name=Snb2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RX NUCLEOTIDE SEQUENCE.
 RP MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."
 RL Meth. Enzymol. 303:19-44(1999).
 [2]
 NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=spinal ganglion;
 RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Izawa M., Fukunishi Y., Komoto H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa T.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barah G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guetlich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 [3]
 NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=spinal ganglion;
 RC THE FANTOM Consortium,
 RA The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 [4]
 NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=spinal ganglion;
 RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,
 RA Komoto H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes."
 RL Genome Res. 10:1617-1630(2000).
 [5]
 NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=spinal ganglion;
 RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki K., Carninci P.,
 RA Komoto H., Akiyama J., Nishi K., Kitanai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara K., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer."
 RL Genome Res. 10:1757-1771(2000).
 [6]
 NUCLEOTIDE SEQUENCE.
 RP STRAIN=C57BL/6J; TISSUE=spinal ganglion;
 RC Adachi J., Aizawa K., Akimura T., Arikawa T., Bono H., Carninci P.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Komoto H., Koyama S.,
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC 1. SIMILARITY: Contains 1 PH domain.
 DR EMBL; AK051630; BAC34698.1; -, mRNA.
 DR HSSP; Q62261; IBTN.
 DR SMR; O8B035; 333-438.
 DR MGI; MGI:98388; Sbn2.
 DR GO; GO:0005634; C:plasma membrane; IDA.
 DR GO; GO:0005886; C:plasma membrane; IDA.
 DR GO; GO:0005518; F:protein binding; IPT.
 DR GO; GO:0007182; P:common-partner SMAD protein phosphorylation; IDA.
 DR GO; GO:0007184; P:SMAD protein nuclear translocation; IDA.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR01993; PH_type.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR001605; Spectrin_PH.
 DR Pfam; PF00169; PH; 1.

DR Pfam; PF00435; Spectrin; 3.
 DR PRINTS; SM00683; SPECTRINPH.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00150; SPEC; 2.
 DR PROSITE; PS50003; PH DOMAIN; 1.
 FT NON_TER 1
 SQ SEQUENCE 497 AA; 56537 MW; BBBA4D95709BF5F0 CRC64;

Query March 28.2%; Score 62; DB 2; Length 497;
 Best Local Similarity 100.0%; Pred. No. 8.6e-55;
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 131 RLTTLELEVARQOEERERKRRPPSPDNTKYSEAESEQMDTSKQDVSONGLPAEQGS 190
 Db 212 RLTTLELEVARQOEERERKRRPPSPDNTKYSEAESEQMDTSKQDVSONGLPAEQGS 271
 Qy 191 PR 192
 Db 272 PR 273

RESULT 18
 ID Q8VI65_RAT PRELIMINARY; PRT; 66 AA.
 AC Q8VI65;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Beta II spectrin-short isoform (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Cai T., Yu P.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF218849; AA556652.1; mRNA.
 DR EMBL; ENSRN00000005434; Rattus norvegicus.
 FT NON_TER 1
 SQ SEQUENCE 66 AA; 7408 MW; CEAC8D69417E3B0 CRC64;

Query March 27.3%; Score 60; DB 2; Length 66;
 Best Local Similarity 100.0%; Pred. No. 1.4e-53;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 TKVSEAESEQMDTSKQDVSONGLPAEQSPRVSQTYQYKNFSRRTASDHSWG 219
 Db 6 TKVSEAESEQMDTSKQDVSONGLPAEQSPRVSQTYQYKNFSRRTASDHSWG 65

RESULT 19
 ID Q9RIY9_MOUSE PRELIMINARY; PRT; 48 AA.
 AC Q9RIY9;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE ELF3 (Fragment).
 GN Name=ELF3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Cai T., Mishra L.;
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF047686; AAD15747.1; Genomic_DNA.
 FT NON_TER 1

FT NON_TER 48
 SQ SEQUENCE 48 AA; 5393 MW; BD9CFB2CAD1632A CRC64;

Query March 21.8%; Score 48; DB 2; Length 48;
 Best Local Similarity 100.0%; Pred. No. 2.6e-41;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 EEBERKRRPPSPDNTKYSEAESEQMDTSKQDVSONGLPAEQGS 192
 Db 1 EEBERKRRPPSPDNTKYSEAESEQMDTSKQDVSONGLPAEQGS 48

RESULT 20
 ID Q4SR6_TETNG PRELIMINARY; PRT; 2275 AA.
 AC Q4SR6;
 DT 13-SEP-2005 (TREMBlrel. 31, Created)
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
 DE Chromosome undetermined SCAF14494, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTENG00013771001;
 OS Tetradon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 CX NCBI_TaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
 RA Mucelli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Doszat C., Segurens B.,
 RA Dailiva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Bilemont C., Skalli Z., Catcicollo L., Poulain J., De Bernardis V.,
 RA Cruaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellie W., Wolff U., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Lauder V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype";
 RL Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG GenomeScope, Whitehead Institute Centre for Genome Research;
 RG Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL; CAAB01014494; CAF96656.1; Genomic_DNA.
 FT NON_TER 1
 SQ SEQUENCE 2275 AA; 261638 MW; 51646A21509AD94 CRC64;

Query March 21.4%; Score 47; DB 2; Length 2275;
 Best Local Similarity 100.0%; Pred. No. 1.2e-38;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LPKPTGRMRHICLENVDKALQFLKQRYVLEKMGSHDIVDGNHRLT 133
 Db 219 LPKPTGRMRHICLENVDKALQFLKQRYVLEKMGSHDIVDGNHRLT 265

RESULT 21
 ID Q4SVU8_TETNG PRELIMINARY; PRT; 2307 AA.
 AC Q4SVU8;
 DT 13-SEP-2005 (TREMBlrel. 31, Created)
 DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
 DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
 DE Chromosome undetermined SCAF13839, whole genome shotgun sequence.

```

DE (Fragment).
GN ORFNames=GSTENG00012286001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Neuteleostomi;
AC Actinopterygii; Neopterygii; Teleostei; Euteleostomi;
OC Acanthomorphi; Acanthopterygii; Perciformes; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Tallon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Litalien G., Dossat C., Segurens B.,
RA Daellva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anhouard V., Jubin C., Castel V., Katinka M., Vacherie B.,
RA Bismont C., Skallit Z., Catelico L., Poulain J., De Bernardis V.,
RA Cruaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Boeak S.,
RA Kellie M., Volff J.N., Gilgo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Landet V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CATION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; CAZ01013839; CAP95574.1; -; Genomic_DNA.
DR InterPro; IPR001589; Actnin_actin_bd.
DR InterPro; IPR001715; Calponin_act_bd.
DR InterPro; IPR001849; PH.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001605; Spectrin_PH.
DR Pfam; PF00307; CH_2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00435; Spectrin; 13.
DR PRINTS; PR00683; SPECTRINPH.
DR SMART; SM00233; PH; 1.
DR SMART; SM00150; SPEC; 14.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH_2.
DR PROSITE; PS50003; PH_DOMAIN; 1.
FT NON_TER 1
FT NON_TER 2307
SQ SEQUENCE 2307 AA; 262847 MW; 4B50PF1DCC68A638 CRC64;

Query Match 20.9%; Score 46; DB 2; Length 2307;
Best Local Similarity 100.0%; Pred. No. 1.3e-37;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 88 PKPTGRMRHICLENVDKALQFLKQRYVLENNWGHSHDYDGNHRT 133
DB 140 PKPTGRMRHICLENVDKALQFLKQRYVLENNWGHSHDYDGNHRLT 185

```

```

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20515613; PubMed=11060238;
RA Liao B.C., Paw B.H., Peters U.L., Zapata A., Pratt S.J., Do C.P.,
RA Lieschke G., Zon L.I.;
RT "Hereditary spherocytosis in zebrafish rising illustrates evolution
RT of erythroid beta-spectrin structure, and function in red cell
RT morphogenesis and membrane stability."
RL Development 127:5123-5132(2000).
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AF262336; AAC03012.1; -; mRNA.
DR HSSP; Q01082; 1BKR.
DR Ensembl; ENSDARG0000030490; Danio rerio.
DR ZFIN; ZDB-GENE-000906-1; spbl.
DR GO; GO:0003779; P:actin binding; IEA.
DR InterPro; IPR001589; Actnin_actin_bd.
DR InterPro; IPR001715; Calponin_act_bd.
DR InterPro; IPR011993; PH_type.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001605; Spectrin_PH.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00435; Spectrin; 17.
DR PRINTS; PR00683; SPECTRINPH.
DR SMART; SM00033; CH; 2.
DR SMART; SM00233; PH; 1.
DR SMART; SM00150; SPEC; 17.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Actin capping; Actin-binding; Cytoskeleton.
SQ SEQUENCE 2357 AA; 272828 MW; 989415B3AD9A71B CRC64;

Query Match 20.9%; Score 46; DB 2; Length 2357;
Best Local Similarity 100.0%; Pred. No. 1.3e-37;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 68 DLRDGRMLIKLEVLGSRLLPKPTGRMRHICLENVDKALQFLKEQ 113
DB 84 DLRDGRMLIKLEVLGSRLLPKPTGRMRHICLENVDKALQFLKEQ 129

RESULT 23
SPTBL_MOUSE STANDARD; PRT; 2127 AA.
ID SPTBL_MOUSE
AC P15508;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DR Spectrum beta chain, erythrocyte (Beta-1 spectrin).
GN Name=Spbl; Synonyms=Spnb-1, Spbl, Spbl1, Spbl2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Eumarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94033578; PubMed=8219239;
RA Bloom M.L., Birkenmeier C.S., Barker J.E.;
RT "Complete nucleotide sequence of the murine erythroid beta-spectrin
RT cDNA and tissue-specific expression in normal and jaundiced mice."
RL Blood 82:2906-2914(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 1054-1289.
RX MEDLINE=88001019; PubMed=3307952;
RA Cioe L., Laurilla P., Mo P., Krebs K., Goodman S., Curtis P.J.;
RT "Cloning and nucleotide sequence of a mouse erythrocyte beta-spectrin

```


RT CDNA." ;
 RL Blood 70:915-920(1987).
 CC -1- FUNCTION: Spectrin is the major constituent of the cytoskeletal
 CC network underlying the erythrocyte plasma membrane. It associates
 CC with band 4.1 and actin to form the cytoskeletal superstructure of
 CC the erythrocyte plasma membrane.
 CC -1- SUBUNIT: Composed of nonhomologous chains, alpha and beta, which
 CC aggregate to form dimers, tetramers, and higher polymers.
 CC -1- MISCELLANEOUS: This complex is anchored to the cytoplasmic face of
 CC the plasma membrane via another protein, ankyrin, which binds to
 CC beta-spectrin and mediates the binding of the whole complex to a
 CC transmembrane protein band 3. The interaction of erythrocyte
 CC spectrin with other proteins through specific binding domains lead
 CC to the formation of an extensive subplasmalemmal meshwork which is
 CC thought to be responsible for the maintenance of the biconcave
 CC shape of human erythrocytes, for the regulation of plasma membrane
 CC components and for the maintenance of the lipid asymmetry of the
 CC plasma membrane.
 CC -1- SIMILARITY: Belongs to the spectrin family.
 CC -1- SIMILARITY: Contains 2 CH (calponin-homology) domains.
 CC -1- SIMILARITY: Contains 17 spectrin repeats.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: S66283; AAB28600.1; -; mRNA.
 CC EMBL: M18641; AAA40126.1; -; mRNA.
 CC HSSP: Q01082; 1BRK.
 CC SMR: P15508; 1055-1264.
 CC DR Ensembl: ENSMUSG0000021061; Mus musculus.
 CC MGI: MGI:98387; Spbn1.
 CC InterPro: IPR001589; Actinin actin bd.
 CC InterPro: IPR001715; Calponin act_bd.
 CC InterPro: IPR002017; Spectrin.
 CC Pfam: PF00307; CH; 2.
 CC DR Pfam: PF00435; Spectrin; 17.
 CC SMART: SM00033; CH; 2.
 CC DR SMART: SM00150; SPEC; 17.
 CC DR PROSITE: PS00019; ACTININ 1; 1.
 CC DR PROSITE: PS00020; ACTININ 2; 1.
 CC DR PROSITE: PS50021; CH; 2.
 CC KW Actin capring; Actin-binding; Cytoskeleton; Erythrocyte; Membrane;
 KW Repeat.
 FT INIT MET 0 0 By similarity.
 FT DOMAIN 1 274 Actin-binding.
 FT DOMAIN 53 157 CH 1.
 FT DOMAIN 172 274 CH 2.
 FT REPEAT 275 383 Spectrin 1.
 FT REPEAT 384 489 Spectrin 2.
 FT REPEAT 490 598 Spectrin 3.
 FT REPEAT 599 704 Spectrin 4.
 FT REPEAT 705 809 Spectrin 5.
 FT REPEAT 810 915 Spectrin 6.
 FT REPEAT 916 1022 Spectrin 7.
 FT REPEAT 1023 1129 Spectrin 8.
 FT REPEAT 1130 1235 Spectrin 9.
 FT REPEAT 1236 1340 Spectrin 10.
 FT REPEAT 1341 1446 Spectrin 11.
 FT REPEAT 1447 1545 Spectrin 12.
 FT REPEAT 1546 1651 Spectrin 13.
 FT REPEAT 1652 1758 Spectrin 14.
 FT REPEAT 1759 1863 Spectrin 15.
 FT REPEAT 1864 1969 Spectrin 16.
 FT REPEAT 1970 2072 Spectrin 17.
 FT SEQUENCE 2127 AA; 245119 MW; 2026FBBD6F775ED1 CRC64;
 Query Match 17.3%; Score 38; DB 1; Length 2127;
 Best Local Similarity 100.0%; Pred. No. 2.2e-29;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 95 MRICLENVDKALQFLKEQVHLENMGSHDIVDGNHRL 132
 |||||
 Db 107 MRICLENVDKALQFLKEQVHLENMGSHDIVDGNHRL 144
 |||||
 RESULT 24
 SPTBL HUMAN STANDARD; PRT; 2136 AA.
 ID SPTBL HUMAN 015510; 015519;
 AC P11277; 015510; 015519;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-FEB-2005 (Rel. 46, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Spectrin beta chain, erythrocyte (Beta-I spectrin).
 GN Name=SPTB; Synonyms=SPTBL;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 OC NCBI_TaxID=9606;
 OX [1]
 RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS 1 AND 3).
 RX MEDLINE=90307707; PubMed=2195026;
 RA Winkelman J.C., Chang J.G., Tse W.T., Scarpa A.L., Marchesi V.T.,
 RT Forget B.G.;
 RL "Full-length sequence of the cDNA for human erythroid beta-spectrin.";
 RL J. Biol. Chem. 265:11827-11832(1990).
 [2]
 RN NUCLEOTIDE SEQUENCE [MRNA] OF 1054-2136 (ISOFORM 2).
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=91056094; PubMed=2243099;
 RA Winkelman J.C., Costa F.F., Linzie B.L., Forget B.G.;
 RT "Beta spectrin in human skeletal muscle. Tissue-specific differential
 RT processing of 3' beta spectrin pre-mRNA generates a beta spectrin
 RT isoform with a unique carboxyl terminus.";
 RL J. Biol. Chem. 265:20449-20454(1990).
 [3]
 RN NUCLEOTIDE SEQUENCE [MRNA] OF 2001-2136.
 RX MEDLINE=91332035; PubMed=1840591;
 RA Gallagher P.G., Tse W.T., Costa F., Scarpa A., Bolvin P., Delaunay J.,
 RT Forget B.G.;
 RT "A splice site mutation of the beta-spectrin gene causing exon
 RT skipping in hereditary elliptocytosis associated with a truncated
 RT beta-spectrin chain.";
 RL J. Biol. Chem. 266:15154-15159(1991).
 [4]
 RN NUCLEOTIDE SEQUENCE [MRNA] OF 927-1754.
 RX MEDLINE=91007291; PubMed=1976574; DOI=10.1016/0378-1119(90)90104-Y;
 RA Yoon S.H., Kentros C.G., Prchal J.T.;
 RT "Identification of an unusual deletion within homologous repeats of
 RT human reticulocyte beta-spectrin and probable peptide polymorphism.";
 RL Gene 91:297-302(1990).
 [5]
 RN NUCLEOTIDE SEQUENCE [MRNA] OF 1333-1431 AND 1908-2136.
 RX MEDLINE=88269838; PubMed=3390609;
 RA Winkelman J.C., Iero T.L., Watkins P.C., Eddy R., Shows T.B.,
 RT Linnebach A.J., Sahr K.E., Kathuria N., Marchesi V.T., Forget B.G.;
 RT "Molecular cloning of the cDNA for human erythrocyte beta-spectrin.";
 RL Blood 72:328-334(1988).
 [6]
 RN NUCLEOTIDE SEQUENCE [MRNA] OF 1208-1481.
 RX MEDLINE=88041127; PubMed=3478706;
 RA Prchal J.T., Morley B.J., Yoon S.-H., Coetzer T.L., Palek J.,
 RT Conboy J.G., Kan Y.W.;
 RT "Isolation and characterization of cDNA clones for human erythrocyte
 RT beta-spectrin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:7468-7472(1987).
 [7]
 RN PROTEIN SEQUENCE OF 1-17.
 RC TISSUE=Platelet;
 RX MEDLINE=22608298; PubMed=12665801; DOI=10.1038/nbt810;
 RA Gevaert K., Goethals M., Martens L., Van Damme J., Staes A.,
 RA Thomas G.R., Vandekerckhove J.;
 RT "Exploring proteomes and analyzing protein processing by mass

RT spectrometric identification of sorted N-terminal peptides.";
RL Nat. Biotechnol. 21:566-569(2003).
RN [8]
RP DOMAINS:
RX MEDLINE=84295638; PubMed=6472478;
RA Speicher D.W., Marchesi V.T.;
RT "Erythrocyte spectrin is comprised of many homologous triple helical
segments.";
RL Nature 311:177-180(1984).
RN [9]
RP REVIEW ON VARIANTS:
RX MEDLINE=97001215; PubMed=8844207;
RX DOI=10.1002/(SICI)1098-1004(1996)8:2<97::AID-HBMU1>3.3.CO;2-W;
RA Maillet P., Allouais N., Mople L., Delaunay J.;
RT "Spectrin mutations in hereditary elliptocytosis and hereditary
spherocytosis.";
RL Hum. Mutat. 8:97-107(1996).
RN [10]
RP VARIANT HE CAGLIARI GLY-2017.
RX MEDLINE=94043025; PubMed=8226774;
RA Sahr K.E., Coetzer T.L., Moy L.S., Derick L.H., Chishti A.H.,
RJ Jacolin P., Lorenzo F., del Giudice E.M., Iolascon A., Gallanello R.,
RA Cao A., Delaunay J., Liu S.-C., Palek J.,
RT "Spectrin Caglieri: an Ala->Gly substitution in helix 1 of beta
spectrin repeat 17 that severely disrupts the structure and self-
association of the erythrocyte spectrin heterodimer.";
RL J. Biol. Chem. 268:22656-22662(1993).
RN [11]
RP VARIANT HS KISSIMMEE ARG-201.
RX MEDLINE=93352802; PubMed=8102379;
RA Becker P.S., Tse W.T., Lux S.B., Forget B.G.;
RT "Beta Spectrin in Kissimmee: a spectrin variant associated with autosomal
dominant hereditary spherocytosis and defective binding to protein
4.1.";
RL J. Clin. Invest. 92:612-616(1993).
RN [12]
RP VARIANT HE PROVIDENCE PRO-2018.
RX MEDLINE=95190014; PubMed=7883966;
RA Gallagher P.G., Weed S.A., Tse W.T., Benoit L., Morrow J.S.,
RJ Marchesi S.L., Mohandas N., Forget B.G.;
RT "Recurrent fatal hydrops fetalis associated with a nucleotide
substitution in the erythrocyte beta-spectrin gene.";
RL J. Clin. Invest. 95:1174-1182(1995).
RN [13]
RP VARIANTS HE VAL-2022 AND ARG-2023.
RX MEDLINE=94289716; PubMed=8018926;
RA Parquet N., Devaux I., Boulanger L., Galand C., Boivin P.,
RJ Lecointe M.-C., Dherny D., Garbarz M.;
RT "Identification of three novel spectrin alpha I/74 mutations in
hereditary elliptocytosis: further support for a triple-stranded
folding unit model of the spectrin heterodimer contact site.";
RL Blood 84:303-308(1994).
RN [14]
RP VARIANT HE PRO-2052.
RX MEDLINE=90369011; PubMed=1975598;
RA Tse W.T., Lecointe M.-C., Costa F.F., Garbarz M., Feo C., Boivin P.,
RJ Dherny D., Forget B.G.;
RT "Point mutation in the beta-spectrin gene associated with alpha I/74
hereditary elliptocytosis. Implications for the mechanism of spectrin
dimer self-association.";
RL J. Clin. Invest. 86:909-916(1990).
RN [15]
RP FUNCTION: Spectrin is the major constituent of the cytoskeletal
network underlying the erythrocyte plasma membrane. It associates
with band 4.1 and actin to form the cytoskeletal superstructure of
the erythrocyte plasma membrane.
CC -1- SUBUNIT: Composed of nonhomologous chains, alpha and beta, which
aggregate to form dimers, tetramers, and higher polymers.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=P11277-1; Sequence=Displayed;
CC Name=2; Synonym=Muscle-specific;
CC IsoId=P11277-2; Sequence=VSP_000719;

CC Name=3;
CC IsoId=P11277-3; Sequence=VSP_007242;
CC -1- DISEASE: Defects in SPTB are a cause of rhesus-unlinked hereditary
elliptocytosis (HE) [MIM:182870, 130600, 166900]; also called
elliptocytosis 2 (EL2). Hereditary elliptocytosis (also known as
hereditary ovalocytosis) is a genetically heterogeneous, autosomal
dominant hematology disorder. It is characterized by variable
hemolytic anemia and elliptical or oval red cell shape. Several
abnormalities in the membrane skeleton have been identified,
including a number that have been localized to spectrin.
CC Elliptocytosis due to defects in SPTB can also be referred to as
elliptocytosis 3 (EL3).
CC -1- DISEASE: Defects in SPTB are the cause of spherocytosis type I
(SPTI) [MIM:182870]. SPTI is a disorder characterized by severe
hemolytic anemia. Inheritance is autosomal dominant.
CC MISCELLANEOUS: This complex is anchored to the cytoplasmic face of
the plasma membrane via another protein, ankyrin, which binds to
beta spectrin and mediates the binding of the whole complex to a
transmembrane protein band 3. The interaction of erythrocyte
spectrin with other proteins through specific binding domains lead
to the formation of an extensive subplasmalemmal network which is
thought to be responsible for the maintenance of the biconcave
shape of human erythrocytes, for the regulation of plasma membrane
components and for the maintenance of the lipid asymmetry of the
plasma membrane.
CC -1- SIMILARITY: Belongs to the spectrin family.
CC -1- SIMILARITY: Contains 2 CH (calponin-homology) domains.
CC -1- SIMILARITY: Contains 17 spectrin repeats.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
CC EMBL: J05500; AAA60578.1; -; mRNA.
CC EMBL: J05500; AAA60579.1; -; mRNA.
CC EMBL: M37884; AAA63259.1; -; mRNA.
CC EMBL: M37885; AAA60571.1; -; mRNA.
CC EMBL: M57948; -; NOT ANNOTATED CDS; mRNA.
CC EMBL: X59510; CAA42057.1; -; mRNA.
CC EMBL: X59511; CAA42098.1; -; mRNA.
CC EMBL: M18054; AAA60572.1; -; mRNA.
CC PIR: A37064; STHB.
CC PDB: 1S35; X-ray; A=1063-1274.
CC SWISS-2DPAGE: P11277; HUMAN.
CC Ensembl: ENSG0000070182; Homo sapiens.
CC HGNC: HGNC:11274; SPTB.
CC MIM: 182870; -.
CC MIM: 130600; -.
CC MIM: 166900; -.
CC GO: GO:0008091; C:spectrin; TAS.
CC GO: GO:0003779; F:actin binding; TAS.
CC InterPro: IPR001589; Actinin_bdn.
CC InterPro: IPR001715; Calponin_act_bdn.
CC Pfam: PF00307; CH; 2.
CC Pfam: PF00435; Spectrin; 17.
CC SMART: SM00033; CH; 2.
CC SMART: SM00150; SPTB; 17.
CC PROSITE: PS00019; ACTININ_1; 1.
CC -----
CC Query Match 17.3%; Score 38; DB 1; Length 2136;
CC Best Local Similarity 100.0%; Pred. No. 2.2e-29;
CC Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
CC QY 95 MRHCLENVKALQFLKQORVHLENGSHDIYDGNHRL 132
CC |||||
CC Db 107 MRHCLENVKALQFLKQORVHLENGSHDIYDGNHRL 144
CC -----
CC RESULT 25
CC Q6XDAO_RAT

```

ID Q6XDA0 RAT PRELIMINARY; PRT; 2137 AA.
AC Q6XDA0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Erythroid spectrin beta.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Chen Y., Huang C.-H.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF283343; AAQ02379.1; -, mRNA.
DR HSSP; Q01082; 1AA2.
DR SMR; Q6XDA0; 1064-1273.
DR GO; GO:0005856; Cytoskeleton; IEA.
DR GO; GO:0003779; F-actin binding; IEA.
DR InterPro; IPR01589; Actbind_actinin.
DR InterPro; IPR001715; Calponin-like.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00435; Spectrin; 17.
DR SMART; SM00033; CH; 2.
DR SMART; SM00150; SPEC; 17.
DR PROSITE; PS00019; ACTININ_1, 1.
DR PROSITE; PS00020; ACTININ_2, 1.
DR PROSITE; PS50021; CH; 2.
SQ SEQUENCE 2137 AA; 246327 MW; 0BC476DF970D878 CRC64;

Query Match 17.3%; Score 38; DB 2; Length 2137;
Best Local Similarity 100.0%; Pred. No. 2.2e-29;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 MRHCLENVKALQFLKEQVHLNMGSHDIVDGNHRL 132
DB 108 MRHCLENVKALQFLKEQVHLNMGSHDIVDGNHRL 145

RESULT 26
Q59FP5 HUMAN PRELIMINARY; PRT; 2332 AA.
AC Q59FP5;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Spectrin, beta, erythrocytic (includes spherocytosis, clinical type I)
DE variant (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Totsuki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
RA Ohara O., Nagase T., Kikuno F.R.;
RT "None Title.";
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AB209415; BAD92652.1; -, mRNA.
DR SMR; Q59FP5; 1068-1277.
DR GO; GO:0003779; F-actin binding; IEA.
DR InterPro; IPR01589; Actinin_actin_bd.
DR InterPro; IPR001715; Calponin_act_bd.
DR InterPro; IPR001849; PH.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001605; Spectrin_PH.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00169; PH; 1.

```

```

DR Pfam; PF00435; Spectrin; 17.
DR PRINTS; PRO0683; SPECTRINPH.
DR SMART; SM00033; CH; 2.
DR SMART; SM00233; PH; 1.
DR SMART; SM00150; SPEC; 17.
DR PROSITE; PS00019; ACTININ_1, 1.
DR PROSITE; PS00020; ACTININ_2, 1.
DR PROSITE; PS50021; CH; 2.
DR PROSITE; PS50003; PH DOMAIN; 1.
KM Actin capping; Actin-binding.
FT NON_TER 1
SQ SEQUENCE 2332 AA; 268239 MW; F133F55BBB8B8AB CRC64;

Query Match 17.3%; Score 38; DB 2; Length 2332;
Best Local Similarity 100.0%; Pred. No. 2.4e-29;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 MRHCLENVKALQFLKEQVHLNMGSHDIVDGNHRL 132
DB 112 MRHCLENVKALQFLKEQVHLNMGSHDIVDGNHRL 149

RESULT 27
Q8WYB3 HUMAN PRELIMINARY; PRT; 67 AA.
AC Q8WYB3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Beta II spectrin-short isoform (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Cai T., Yu P.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF218848; AAL56651.1; -, mRNA.
FT NON_TER 1
SQ SEQUENCE 67 AA; 7551 MW; F832D21E1BDC5DD CRC64;

Query Match 16.8%; Score 37; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 8.2e-30;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 QVSQNGLPALQGSPPVSYRSQTYQYKNFNSRRRTASD 214
DB 25 QVSQNGLPALQGSPPVSYRSQTYQYKNFNSRRRTASD 61

RESULT 28
Q9WUX0 RAT PRELIMINARY; PRT; 82 AA.
AC Q9WUX0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Beta II spectrin (Fragment).
GN Name=Spnbl; Synonyma=sptbnl;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA TISSUE=heart;
RX MEDLINE=20267884; PubMed=10806113;
RA Hayes N.V.L., Scott C., Heerikens E., Ohanian V., Maggs A.M.,
RA Pinder J.C., Kordeli E., Baines A.J.;
RT "Identification of a novel C-terminal variant of betaII spectrin: two

```

RT isoforms of beta1 spectrin have distinct intracellular locations and activities."

RL J. Cell Sci. 113:2023-2034(2000).

DR EMBL; AJ242018; CAB42616.1; -, mRNA.

DR RGD; 621726; Spnbl.

FT NON_TER 1

SQ SEQUENCE 82 AA; 9526 MW; 6FDA86568647C6B9 CRC64;

Query Match 16.8%; Score 37; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 9.9e-30;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 QVSQNGLPABQSGPRYSRQTYQNYKNFNSRTASD 214
DB 40 QVSQNGLPABQSGPRYSRQTYQNYKNFNSRTASD 76

RESULT 29
QSD002_RAT PRELIMINARY; PRT; 516 AA.

AC QSD002;
DT 10-MAY-2005 (T-EMBLrel. 30, Created)
DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
DE 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
DE Spnbl protein (Fragment).
GN Name=Spnbl2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_Taxid=10116;
[1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22388237; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schaller G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Ditchenko L., Marinina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.D., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Ovary;
RG NIH MGC Project;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; BC090340; AAH90340.1; -, mRNA.
DR SMR; QSD002; 352-457.
DR InterPro; IPR001849; PH.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001605; Spectrin_PH.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00435; Spectrin; 2.
DR PRINTS; PR00683; SPECTRINPH.
DR SMART; SM00233; PH; 1.
DR SMART; SM00150; SPEC; 2.
DR PROSITE; PS50003; PH_DOMAIN; 1.
FT NON_TER 1

SQ SEQUENCE 516 AA; 58595 MW; 2969B75C89F71A3A CRC64;

Query Match 14.5%; Score 32; DB 2; Length 516;
Best Local Similarity 100.0%; Pred. No. 8.6e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KYSEEAESQMDTSKDDQVSQNGLPABQSGPR 192
DB 261 KYSEEAESQMDTSKDDQVSQNGLPABQSGPR 292

RESULT 30
Q4SF10_TETNG PRELIMINARY; PRT; 2413 AA.

AC Q4SF10;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Chromosome 7 SCAP14601, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00019086001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_Taxid=99863;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozou-Costas C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Rubin C., Casteille V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Casteille L., Poulain J., De Berardinis V.,
RA Crnaud C., Duprat S., Broctier P., Coutancieu J.P., Gonzy J.,
RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Queller F., Sautin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissbach J., Roest Crollins H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype."

RT Nature 431:946-957(2004).
[2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -1- FUNCTION: Plays a key role in the control of the eukaryotic cell
CC cycle. It is required in higher cells for entry into S-phase and
CC mitosis. Component of the kinase complex that phosphorylates the
CC repetitive C-terminus of RNA polymerase II. Catalytic component of
CC MPF (By similarity).
CC -1- SUBUNIT: Forms a stable but non-covalent complex with cyclin B in
CC mature oocytes (By similarity).
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; CA601014601; CAG00602.1; -, Genomic DNA.
DR InterPro; IPR001589; Actinin_actin_bd.
DR InterPro; IPR001715; Actinin_act_bd.
DR InterPro; IPR001649; PH.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001605; Spectrin_PH.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00435; Spectrin; 17.
DR PRINTS; PR00683; SPECTRINPH.
DR SMART; SM00033; CH; 2.
DR SMART; SM00233; PH; 1.
DR SMART; SM00150; SPEC; 17.
DR PROSITE; PS00019; ACTININ_1; 1.

DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
DR PROSITE; PS50003; PH DOMAIN; 1.
FT NON TER 1
SQ SEQUENCE 2413 AA; 277305 MW; 8EA3A927A2C17FBB CRC64;

Query Match 12.7%; Score 28; DB 2; Length 2413;
Best Local Similarity 100.0%; Pred. No. 5,2e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 28 LEGRFQLODEREAVOKTFTKVNSHL 55
Db 79 LEGRFQLODEREAVOKTFTKVNSHL 106

RESULT 31
O5XG36_XENLA PRELIMINARY; PRT; 794 AA.
ID O5XG36_XENLA PRELIMINARY; PRT; 794 AA.
AC O5XG36;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DE LOC398511 protein (Fragment).
GN Name=LOC398511;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RT Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RT Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC094634; AA04634.1; -, mRNA.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR001589; Actbind_actin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR02017; Spectrin.
DR Pfam; PF00307; CH; 2.

DR Pfam; PF00435; Spectrin; 5.
DR SMART; SM00033; CH; 2.
DR SMART; SM00150; SPEC; 4.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
FT NON TER 794
SQ SEQUENCE 794 AA; 92657 MW; 96F2259167965456 CRC64;

Query Match 12.3%; Score 27; DB 2; Length 794;
Best Local Similarity 100.0%; Pred. No. 1.9e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 87 LPKPTGMRHICLENVDKALQFLKEQ 113
Db 103 LPKPTGMRHICLENVDKALQFLKEQ 129

RESULT 32
O7ZX11_XENLA PRELIMINARY; PRT; 1060 AA.
ID O7ZX11_XENLA PRELIMINARY; PRT; 1060 AA.
AC O7ZX11;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DE LOC398511 protein (Fragment).
GN Name=LOC398511;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RT Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RT Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046267; AA046267.1; -, mRNA.
DR HSP; 001082; 1BR.
DR GO; GO:0003779; F:actin binding; IEA.

```

DR InterPro; IPR001589; Acbind_actnin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00307; CH_2.
DR SMART; SM00435; Spectrin; 7.
DR SMART; SM00435; CH_2.
DR SMART; SM00150; SPEC; 7.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
DR NON_TER 1060 AA; 122756 MW; ABD56DC56C4C11B CRC64;
SQ SEQUENCE 1060 AA; 122756 MW; ABD56DC56C4C11B CRC64;

Query Match 12.3%; Score 27; DB 2; Length 1060;
Best Local Similarity 100.0%; Pred. No. 2.5e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LPKPTGMRHICLENVDKALQFLKEQ 113
Db 103 LPKPTGMRHICLENVDKALQFLKEQ 129

RESULT 33
SFTN2 RAT STANDARD; PRT; 2388 AA.
AC Q9QWN8; O88197; Q9ES68;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Spectrin beta chain, brain 2 (Spectrin, non-erythroid beta chain 2)
DE (Beta-III spectrin) (SPNB-3) (beta Spili sigma 1) (Spectrin-like
DE protein GTPAP41).
GN Name=Spbn2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=98340016; PubMed=9704016; DOI=10.1016/S0169-328X(98)00068-0;
RA Ohara O., Ohara R., Yamakawa H., Nakajima D., Nakayama M.,
RA Sakaguchi G., Orita S., Naito A., Maeda M., Igataishi H., Sasaki T.,
RA Takai Y.,
RA "A novel brain-specific isoform of beta spectrin: isolation and its
RA interaction with Munc13."
RT Biochem. Biophys. Res. Commun. 248:846-851(1998).
RL [3]
RN NUCLEOTIDE SEQUENCE.
RP Rothstein J.D., Jackson M.,
RA "GTPAP41 - a cytoskeletal modulator of EAAT4 glutamate transporter."
RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RL [1]
RC FUNCTION: Probably plays an important role in neuronal membrane
RC skeleton.
CC [-] TISSUE SPECIFICITY: Abundantly transcribed in the brain. Neurons
CC are the predominant cell-type to express the gene. Found
CC abundantly in Purkinje cells.
CC [-] SIMILARITY: Belongs to the spectrin family.
CC [-] SIMILARITY: Contains 2 CH (calponin-homology) domains.
CC [-] SIMILARITY: Contains 1 PH domain.
CC [-] SIMILARITY: Contains 17 spectrin repeats.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB008551; BAA32699.1; -; mRNA.
CC EMBL; AB001347; BAA32473.1; -; mRNA.
CC EMBL; AF225960; AAG28596.1; -; mRNA.
CC PIR; J02711; J02711.
CC HSSP; 001082; 1BRK.
CC DR SMR; Q9QWN8; 2212-2333.
CC Ensembl; ENSRNOCG00000019564; Rattus norvegicus.
CC GO; GO:0005737; C:cytoplasm; TAS.
CC GO; GO:0016363; C:nuclear matrix; TAS.
CC GO; GO:0008091; C:spectrin; TAS.
CC GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
CC GO; GO:0017158; F:regulation of calcium ion-dependent exocytosis; TAS.
CC GO; GO:0016081; P:synaptic vesicle docking during exocytosis; TAS.
CC DR InterPro; IPR001589; Actnin_actn_bd.
CC DR InterPro; IPR001715; Calponin_act_bd.
CC DR InterPro; IPR01849; PH_type.
CC DR InterPro; IPR02017; Spectrin.
CC DR InterPro; IPR01605; Spectrin_PH.
CC Pfam; PF00307; CH_2.
CC Pfam; PF00169; PH_1.
CC Pfam; PF00435; Spectrin; 17.
CC PRINTS; PR00683; SPECTRINPH.
CC SMART; SM00033; CH; 2.
CC SMART; SM00233; PH; 1.
CC SMART; SM00150; SPEC; 17.
CC PROSITE; PS00019; ACTININ_1; 1.
CC PROSITE; PS00020; ACTININ_2; 1.
CC PROSITE; PS50021; CH; 2.
CC DR PROSITE; PS50003; PH_DOMAIN; 1.
CC KW Actin capring; Actin-binding; Cytoskeleton; Membrane; Repeat.
CC FT DOMAIN 1 278 Actin-binding (By similarity).
CC FT DOMAIN 176 278 CH 1.
CC FT REPEAT 305 415 Spectrin 1.
CC FT REPEAT 425 529 Spectrin 2.
CC FT REPEAT 531 639 Spectrin 3.
CC FT REPEAT 641 745 Spectrin 4.
CC FT REPEAT 747 850 Spectrin 5.
CC FT REPEAT 852 956 Spectrin 6.
CC FT REPEAT 958 1063 Spectrin 7.
CC FT REPEAT 1065 1170 Spectrin 8.
CC FT REPEAT 1172 1276 Spectrin 9.
CC FT REPEAT 1278 1381 Spectrin 10.
CC FT REPEAT 1383 1486 Spectrin 11.
CC FT REPEAT 1488 1586 Spectrin 12.
CC FT REPEAT 1588 1692 Spectrin 13.
CC FT REPEAT 1694 1799 Spectrin 14.
CC FT REPEAT 1801 1905 Spectrin 15.
CC FT REPEAT 1907 2011 Spectrin 16.
CC FT REPEAT 2013 2075 Spectrin 17.
CC FT DOMAIN 2218 2328 PH.
CC FT CONFLICT 326 328 VTL -> GTF (in Ref. 3).
CC FT CONFLICT 543 543 L -> F (in Ref. 2).
CC FT CONFLICT 608 608 D -> G (in Ref. 3).
CC FT CONFLICT 867 867 L -> P (in Ref. 3).
CC FT CONFLICT 908 908 V -> I (in Ref. 3).
CC FT CONFLICT 948 948 D -> G (in Ref. 3).
CC FT CONFLICT 1156 1157 EL -> GA (in Ref. 2).
CC FT CONFLICT 1194 1194 F -> V (in Ref. 2 and 3).
CC FT CONFLICT 1555 1555 G -> R (in Ref. 3).
CC FT CONFLICT 1769 1769 R -> W (in Ref. 3).
CC SQ SEQUENCE 2388 AA; 271064 MW; 3BC8966AF0665F19 CRC64;

Query Match 12.3%; Score 27; DB 1; Length 2388;
Best Local Similarity 100.0%; Pred. No. 5.5e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LPKPTGMRHICLENVDKALQFLKEQ 113
|||||

```

DB 103 LPKPTKGRMRHICLENVDKALQFLKEQ 129

RESULT 34
ID 068FG2_MOUSE PRELIMINARY; PRT; 2388 AA.
AC 068FG2;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Spectrin beta 3.
GN Name=Spm3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
CC Muroidea; Muridae; Murinae; Mus.
CC MCB1_TaxID=10090;
CC [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6; TISSUE=Brain;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stopleckon M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Scherch A., Schin J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RG NIH MGC Project;
RL Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
CC EMBL: BC079860; AAH79860.1; -, mRNA.
CC SMR: 068FG2; 2212-2333.
CC Ensembl: ENSMUSG0000067889; Mus musculus.
CC MGI: MGI:1313261; Spm3.
CC GO: GO:0003779; F-actin binding; IEA.
CC InterPro: IPR001589; Actinin actin bd.
CC InterPro: IPR001715; Calponin_act_bd.
CC InterPro: IPR001849; PH.
CC InterPro: IPR011993; PH type.
CC InterPro: IPR02017; Spectrin.
CC InterPro: IPR001605; Spectrin_PH.
CC Pfam: PF00307; CH; 2.
CC Pfam: PF00169; PH; 1.
CC Pfam: PF00435; Spectrin; 17.
CC PRINTS: PRO0663; SPECTRINPH.
CC SMART: SM00033; CH; 2.
CC SMART: SM00233; PH; 1.
CC SMART: SM00150; SPM3; 17.
CC PROSITE: PS00019; ACTININ_1; 1.
CC PROSITE: PS00020; ACTININ_2; 1.
CC PROSITE: PS50023; CH; 2.
CC PROSITE: PS50003; PH DOMAIN; 1.
CC Actin capring; Actin-binding; Cytoskeleton.
CC SEQUENCE 2388 AA; 270923 MW; 19DA7460B75576D1 CRC64;
Query Match 12.3%; Score 27; DB 2; Length 2388;

Best Local Similarity 100.0%; Pred. No. 5,5e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LPKPTKGRMRHICLENVDKALQFLKEQ 113
DB 103 LPKPTKGRMRHICLENVDKALQFLKEQ 129

RESULT 35
ID SPTN2_HUMAN STANDARD; PRT; 2390 AA.
AC 015020; 014872; 014873;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Spectrin beta chain, brain 2 (Spectrin, non-erythroid beta chain 2)
DE (beta-III spectrin).
GN Name=SPTN2; Synonyms=KIA00302;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
CC Homo.
CC NCB1_TaxID=9606;
CC [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
RP TISSUE=Brain;
RC MEDLINE=9734984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150 (1997).
RN [2]
RP NUCLEOTIDE SEQUENCE OF 1-34.
RC TISSUE=Brain;
RX MEDLINE=99045654; PubMed=9826670; DOI=10.1073/pnas.95.24.14158;
RA Stankewich M.C., Tse W.T., Peters L.L., Chng Y., John K.M.,
RA Stachek P.R., Devarejan P., Morrow J.S., Lux S.E.;
RT "A widely expressed betaactin spectrin associated with Golgi and
RT cytoplasmic vesicles.";
RT Proc. Natl. Acad. Sci. U.S.A. 95:14158-14163 (1998).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 1900-2390 (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RA Tse W.T., Peters L.L., John K.M., Lux S.E.;
RT "SPTN2, a new, widely expressed beta III spectrin gene located on
RT human chromosome 11q13 and mouse chromosome 19.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Probably plays an important role in neuronal membrane
CC skeleton.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named Isoforms=2;
CC Name=1;
CC IsoId=015020-1; Sequence=Displayed;
CC Name=2;
CC IsoId=015020-2; Sequence=VSP_000722;
CC -1- TISSUE SPECIFICITY: Highly expressed in brain, kidney, pancreas,
CC and liver, and at lower levels in lung and placenta.
CC -1- SIMILARITY: Belongs to the spectrin family.
CC -1- SIMILARITY: Contains 2 CH (calponin-homology) domains.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 17 spectrin repeats.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AB008567; BAA32700.1; -, mRNA.
CC EMBL: AF079569; AAC80006.1; -, mRNA.
CC EMBL: AF026487; AAC79502.1; -, mRNA.

```

DR EMBL: AF026488; AAC79503.1; -; mRNA.
DR EMBL: AF026488; AAC79504.1; -; mRNA.
DR PDB: 1WJM; NMR; A=2219-2326.
DR SMR: O15020; 2212-2333.
DR Eneemb1; ENSG00000173898; Homo sapiens.
DR HGNC; HGNC:11276; SPTBN2.
DR H-InVDB; HIX0021762; -.
DR MIM; 604985; -.
DR GO; GO:0030138; C:COPII-coated vesicle; TAS.
DR GO; GO:0008093; C:spectrin; IEA.
DR GO; GO:0003779; F:actin binding; TAS.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
DR GO; GO:0016192; P:vesicle-mediated transport; IDA.
DR InterPro; IPR001589; Actinin actin bd.
DR InterPro; IPR001715; Calponin_act_bd.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH_type.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001605; Spectrin_PH.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00435; Spectrin; 17.
DR PRINTS; PR00683; SPECTRINPH.
DR SMART; SM00033; CH; 2.
DR SMART; SM00233; PH; 1.
DR SMART; SM00150; SPEC; 17.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR 3D-structure; Actin capping; Actin-binding; Alternative splicing;
KW Cytoskeleton; Membrane; Repeat.
FT DOMAIN 1 278 Actin-binding (By similarity).
FT DOMAIN 57 161 CH 1.
FT DOMAIN 176 278 CH 2.
FT REPEAT 305 415 Spectrin 1.
FT REPEAT 425 529 Spectrin 2.
FT REPEAT 531 639 Spectrin 3.
FT REPEAT 641 745 Spectrin 4.
FT REPEAT 747 850 Spectrin 5.
FT REPEAT 852 956 Spectrin 6.
FT REPEAT 958 1063 Spectrin 7.
FT REPEAT 1065 1170 Spectrin 8.
FT REPEAT 1172 1276 Spectrin 9.
FT REPEAT 1278 1381 Spectrin 10.
FT REPEAT 1383 1486 Spectrin 11.
FT REPEAT 1488 1586 Spectrin 12.
FT REPEAT 1588 1692 Spectrin 13.
FT REPEAT 1694 1799 Spectrin 14.
FT REPEAT 1801 1905 Spectrin 15.
FT REPEAT 1907 2011 Spectrin 16.
FT REPEAT 2013 2075 Spectrin 17.
FT DOMAIN 2218 2328 PH.
FT VARSPLIC 2314 2330 AEMSSMLRVNNAIAITASASGEPEPVVPTTRGNTDAMT
MPVPSVGAEGVPLRSKGRERERKSPFKQK -> V
SCPSGSLVPRQKLPAADSPFPVLPFLPGVLGKTKGV
RRPHQALPV (in isoform 2).
/FTID=VSP_000722.
SQ SEQUENCE 2390 AA; 271295 MW; E8C3033DD25A674C CRC64;

Query Match 12.3%; Score 27; DB 1; Length 2390;
Best Local Similarity 100.0%; Pred. No. 5.5e-18;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DR 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DS Beta-spectrin (Fragment).
GN Name=HSP71;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RX MEDLINE=9811272; PubMed=9450796;
RX DOI=10.1046/j.1365-2141.1998.00530.x;
RA Garbarr M., Galand C., Bibes D., Bournier O., Devaux I.,
RA Haroubeau J.L., Grandchamp B., Dhery D.,
RT "A 5' splice region G->C mutation in exon 3 of the human beta-
RT spectrin gene leads to decreased levels of beta-spectrin mRNA and is
RT responsible for dominant hereditary spherocytosis (spectrin Guemene-
RT Pentao).";
RL Br. J. Haematol. 100:90-98(1998).
DR EMBL: AF013172; AAC02786.1; -; Genomic_DNA.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR001589; Actbind_actnin.
DR InterPro; IPR001715; Calponin-like.
DR Pfam; PF00307; CH; 1.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS50021; CH; 1.
FT NON_TER 1 1
FT NON_TER 51 51
SQ SEQUENCE 51 AA; 6035 MW; 29F9A04CD7A8031 CRC64;

Query Match 11.4%; Score 25; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 36
Q71VG2_HUMAN
ID Q71VG2_HUMAN PRELIMINARY; PRT; 51 AA.
AC Q71VG2;

```

```

Qy 87 LKPTKGRMRHCLNVDKALQFLKEQ 113
Db 103 LKPTKGRMRHCLNVDKALQFLKEQ 129

```

```

RESULT 37
Q71S06_HUMAN PRELIMINARY; PRT; 2002 AA.
ID Q71S06_HUMAN PRELIMINARY; PRT; 2002 AA.
AC Q71S06;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DS Beta-spectrin IV.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
NCBI_TaxID=9606;
RX MEDLINE=9811272; PubMed=9450796;
RX DOI=10.1046/j.1365-2141.1998.00530.x;
RA Garbarr M., Galand C., Bibes D., Bournier O., Devaux I.,
RA Haroubeau J.L., Grandchamp B., Dhery D.,
RT "A 5' splice region G->C mutation in exon 3 of the human beta-
RT spectrin gene leads to decreased levels of beta-spectrin mRNA and is
RT responsible for dominant hereditary spherocytosis (spectrin Guemene-
RT Pentao).";
RL Br. J. Haematol. 100:90-98(1998).
DR EMBL: AF013172; AAC02786.1; -; Genomic_DNA.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR001589; Actbind_actnin.
DR InterPro; IPR001715; Calponin-like.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00435; Spectrin; 16.
DR SMART; SM00033; CH; 2.
DR SMART; SM00150; SPEC; 14.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
SQ SEQUENCE 2002 AA; 225564 MW; 627E9616D192AE30 CRC64;

```


Query Match 10.0%; Score 22; DB 2; Length 2002;
 Best Local Similarity 100.0%; Pred. No. 6.8e-13;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 DEREAVQKKTFTKWNHSHARV 58
 |||||
 Db 57 DEREAVQKKTFTKWNHSHARV 78

RESULT 38
 Q912B6 MOUSE PRELIMINARY; PRT; 2555 AA.
 AC Q912B6
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-FEB-2005 (TREMBlrel. 29, Last annotation update)
 DE Beta4-spectrin (Beta-spectrin 4).
 GN Name=Spnb4;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21419653; PubMed=11528393; DOI=10.1038/ng710;
 RA Parkinson N.J., Olsson C.L., Hallows J.L., McKee-Johnson J.,
 RA Keogh B.P., Noben-Trauth K., Kujawa S.G., Tempel B.L.;
 RT "Mutant beta-spectrin 4 causes auditory and motor neuropathies in
 RT quivering mice".
 RL Nat. Genet. 29:61-65(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6;
 RA Parkinson N.J., Olsson C.L., Hallows J.L., Kujawa S.G.,
 RA McKee-Johnson J., Noben-Trauth K., Keogh B.P., Tempel B.L.;
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 PH domain.
 DR EMBL; AY032720; AAK49014.1; -; Genomic_DNA.
 DR EMBL; AY032655; AAK38731.1; -; mRNA.
 DR EMBL; AY032702; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032700; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032698; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032694; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032696; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032693; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032697; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032701; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032708; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032709; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032710; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032711; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032712; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032713; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032714; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032705; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032703; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032704; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032715; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032716; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032717; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032718; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032719; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032706; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032707; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032699; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032695; AAK49014.1; JOINED; Genomic DNA.
 DR EMBL; AY032692; AAK49014.1; JOINED; Genomic DNA.
 DR HSSP; Q01082; 1BKR.
 DR MGI; MGI:1890574; Spnb4.
 DR GO; GO:0030423; C:axon; IDA.
 DR GO; GO:0043203; C:axon hillock; IDA.

DR GO; GO:0005737; C:cytoplasm; ISS.
 DR GO; GO:0016363; C:nuclear matrix; ISS.
 DR GO; GO:0016505; C:PM body; ISS.
 DR GO; GO:0008091; C:spectrin; ISS.
 DR GO; GO:0003779; F:actin binding; ISS.
 DR GO; GO:0030506; F:ankyrin binding; ISS.
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; ISS.
 DR GO; GO:0007016; P:cytoskeletal anchoring; ISS.
 DR GO; GO:0007605; P:perception of sound; IMP.
 DR GO; GO:0019226; P:transmission of nerve impulse; IMP.
 DR GO; GO:0016192; P:vesicle-mediated transport; ISS.
 DR InterPro; IPR001589; Actnin_actn_bd.
 DR InterPro; IPR001715; Calponin_act_bd.
 DR InterPro; IPR01849; PH.
 DR InterPro; IPR011993; PH_type.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR001605; Spectrin_PH.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00435; Spectrin; 17.
 DR PRINTS; PR00683; SPECTRINPH.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00150; SPEC; 16.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR Actin capring; Actin-binding; Cytoskeleton.
 DR ACIN capring; Actin-binding; Cytoskeleton.
 SQ SEQUENCE 2555 AA; 288135 MW; 39FEC7E410289D45 CRC64;

Query Match 10.0%; Score 22; DB 2; Length 2555;
 Best Local Similarity 100.0%; Pred. No. 8.6e-13;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 DEREAVQKKTFTKWNHSHARV 58
 |||||
 Db 58 DEREAVQKKTFTKWNHSHARV 79

RESULT 39
 Q912B6 MOUSE PRELIMINARY; PRT; 2561 AA.
 AC Q912B6
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE BetaIV-spectrin signal.
 GN Name=Spnb4;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21666188; PubMed=11807096; DOI=10.1083/jcb.200110003;
 RA Komada M., Soriano P.;
 RT "[beta]IV-spectrin regulates sodium channel clustering through
 RT ankyrin-G at axon initial segments and nodes of Ranvier.";
 RL J. Cell Biol. 156:337-348(2002).
 CC -1- SIMILARITY: Contains 1 PH domain.
 DR EMBL; AB055618; BAB83243.1; -; mRNA.
 DR HSSP; Q01082; 1BKR.
 DR MGI; MGI:1890574; Spnb4.
 DR GO; GO:0030424; C:axon; IDA.
 DR GO; GO:0043203; C:axon hillock; IDA.
 DR GO; GO:0005737; C:cytoplasm; IDA.
 DR GO; GO:0016363; C:nuclear matrix; IDA.
 DR GO; GO:0007605; P:perception of sound; IMP.
 DR GO; GO:0019226; P:transmission of nerve impulse; IMP.
 DR InterPro; IPR001589; Actnin_actn_bd.
 DR InterPro; IPR001715; Calponin_act_bd.

```

DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001605; Spectrin_PH.
DR Pfam; PF00307; CH 2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00435; Spectrin; 17.
DR PRINTS; PR00683; SPECTRINPH.
DR SMART; SM00033; CH; 2.
DR SMART; SM00233; PH; 1.
DR SMART; SM00150; SPC; 16.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
DR PROSITE; PS50003; PH DOMAIN; 1.
KW Actin capping; Actin-binding; Cytoskeleton.
SQ SEQUENCE 2561 AA; 288809 MW; 4DF01A9180140F33 CRC64;

Query Match 10.0%; Score 22; DB 2; Length 2561;
Best Local Similarity 100.0%; Pred.No. 8.6e-13;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 DDERAVQKKTFTKWNSHLARY 58
Db 58 DDERAVQKKTFTKWNSHLARY 79

RESULT 40
SPTN4_HUMAN STANDARD; PRT; 2564 AA.
ID SPTN4_HUMAN STANDARD; PRT; 2564 AA.
AC Q9H254; Q9H1K7; Q9H1K8; Q9H1K9; Q9H3G8; Q9HCD0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DDT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Spectrin beta chain, brain 3 (Spectrin, non-erythroid beta chain 3)
DE (Beta-IV spectrin).
GN Name=SPTN4; Synonyms=KTA1642, SPTBN3;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea;
OC Homo; NCBI_TaxID=9606;
OC NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
RX MEDLINE=21316449; PubMed=11294830; DOI=10.1074/jbc.M009307200;
RA Tse W.T., Tang J., Jin O., Kozegren C., John K.M., Kung A.L.,
RA Gwynn B., Peters J.L., Lux S.E.;
RT "A new spectrin, beta-IV, has a major truncated isoform that
RT associates with promyelocytic leukemia protein nuclear bodies and the
RT nuclear matrix."
RL J. Biol. Chem. 276:23974-23985(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE (ISOFORMS 1; 2; 3 AND 4).
RX MEDLINE=20539976; PubMed=11086001; DOI=10.1083/jcb.151.5.985;
RA Berghs S., Aggujaro D., Ditz R. Jr., Maksimova E., Stebach P.,
RA Herml J.-M., Zhang J.-P., Philbrick W., Slepnev V., Ort T.,
RA Solimena M.;
RT "BetaIV spectrin, a new spectrin localized at axon initial segments
RT and nodes of Ranvier in the central and peripheral nervous system."
RL J. Cell Biol. 151:985-1002(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 386-2382 (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro."
RL DNA Res. 7:273-281(2000).
CC -1- ALTERNATIVE PRODUCTS;
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1;
CC IsoId=Q9H254-1; Sequence=Displayed;

```

```

CC CC Name=2;
CC CC IsoId=Q9H254-2; Sequence=VSP_000725, VSP_000726;
CC CC Name=3;
CC CC IsoId=Q9H254-3; Sequence=VSP_000723, VSP_000724;
CC CC Name=4;
CC CC IsoId=Q9H254-4; Sequence=VSP_000727, VSP_000728;
CC CC -1- TISSUE SPECIFICITY: Abundantly expressed in brain and pancreatic
CC CC islets.
CC CC -1- SIMILARITY: Belongs to the spectrin family.
CC CC -1- SIMILARITY: Contains 2 CH (calponin-homology) domains.
CC CC -1- SIMILARITY: Contains 1 PH domain.
CC CC -1- SIMILARITY: Contains 18 spectrin repeats.
CC CC -----
CC CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use as long as its content is in no way modified and this statement is not
CC CC removed.
CC CC -----
CC CC EMBL; AF311855; AAG42473.1; -; mRNA.
CC CC EMBL; AF082075; AAG38874.1; -; mRNA.
CC CC EMBL; AY004226; AAF93171.1; -; mRNA.
CC CC EMBL; AY004226; AAF93172.1; -; mRNA.
CC CC EMBL; AY004227; AAF93173.1; -; mRNA.
CC CC EMBL; AB046862; BAB13468.1; -; mRNA.
CC CC HSSP; Q01082; 1BRK.
CC CC DR Ensembl; ENSG00000160460; Homo sapiens.
CC CC HGN; HGNC:14896; SPTBN4.
CC CC MIM; 606214; -;
CC CC GO; GO:0005737; C:cytoplasm; IDA.
CC CC GO; GO:0016363; C:nuclear matrix; IDA.
CC CC GO; GO:001605; C:PMI body; IDA.
CC CC GO; GO:0008093; C:spectrin; IDA.
CC CC GO; GO:0003779; F:actin binding; TAS.
CC CC GO; GO:0030506; F:ankyrin binding; IDA.
CC CC GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
CC CC GO; GO:0007016; P:cytoskeletal anchoring; TAS.
CC CC GO; GO:0016192; P:vesicle-mediated transport; TAS.
CC CC DR InterPro; IPR001589; Actinin actin bd.
CC CC DR InterPro; IPR001715; Calponin_act_bd.
CC CC DR InterPro; IPR001849; PH.
CC CC DR InterPro; IPR011993; PH type.
CC CC DR InterPro; IPR002017; Spectrin.
CC CC DR InterPro; IPR001605; Spectrin_PH.
CC CC Pfam; PF00307; CH; 2.
CC CC Pfam; PF00169; PH; 1.
CC CC Pfam; PF00435; Spectrin; 17.
CC CC PRINTS; PR00683; SPECTRINPH.
CC CC SMART; SM00033; CH; 2.
CC CC SMART; SM00233; PH; 1.
CC CC SMART; SM00150; SPC; 16.
CC CC PROSITE; PS00019; ACTININ_1; 1.
CC CC PROSITE; PS00020; ACTININ_2; 1.
CC CC PROSITE; PS50021; CH; 2.
CC CC PROSITE; PS50003; PH DOMAIN; 1.
CC CC KW Actin capping; Actin-binding; Alternative splicing; Cytoskeleton;
CC KW Membrane; Repeat.
CC FT DOMAIN 1 282 Actin-binding (By similarity).
CC FT DOMAIN 1 165 CH 1.
CC FT DOMAIN 180 282 CH 2.
CC FT REPEAT 309 354 Spectrin 1.
CC FT REPEAT 398 419 Spectrin 2.
CC FT REPEAT 429 533 Spectrin 3.
CC FT REPEAT 535 642 Spectrin 4.
CC FT REPEAT 644 771 Spectrin 5.
CC FT REPEAT 773 879 Spectrin 6.
CC FT REPEAT 881 985 Spectrin 7.
CC FT REPEAT 1019 1086 Spectrin 8.
CC FT REPEAT 1088 1197 Spectrin 9.
CC FT REPEAT 1199 1303 Spectrin 10.
CC FT REPEAT 1305 1408 Spectrin 11.
CC FT REPEAT 1410 1513 Spectrin 12.
CC FT REPEAT 1515 1619 Spectrin 13.

```

```

FT REPEAT 1621 1725 Spectrin 14.
FT REPEAT 1727 1832 Spectrin 15.
FT REPEAT 1834 1940 Spectrin 16.
FT REPEAT 1942 2046 Spectrin 17.
FT REPEAT 2048 2107 Spectrin 18.
FT DOMAIN 2418 2527 PH.
FT VARSPLIC 1 1257 Missing (in isoform 3).
FT VARSPLIC 1258 1286 /FTid=VSP_000723.
FT VARSPLIC 1287 1309 AVOAEGILRGGINYGEAOAVRLLLEK -> MPHYPSCS
FT VARSPLIC 1309 /FTid=VSP_000724. SAPIGTPIPTIOIQLEKRR (in isoform 3).
FT VARSPLIC 1310 2564 /FTid=VSP_000725. PPIYPRSSS (in isoform 2).
FT VARSPLIC 2113 2154 /FTid=VSP_000726. Missing (in isoform 2).
FT VARSPLIC 2155 2564 /FTid=VSP_000728. Missing (in isoform 4).
FT CONFLICT 604 608 /FTid=VSP_000728. Missing (in Ref. 2).
FT CONFLICT 714 714 L -> S (in Ref. 2).
FT CONFLICT 1189 1189 E -> K (in Ref. 2).
FT CONFLICT 1193 1193 E -> K (in Ref. 2).
FT CONFLICT 1331 1331 G -> S (in Ref. 1).
SQ SEQUENCE 2564 AA; 288985 MW; 52CDE7D1DB01BCC CRC64;

```

Query Match 10.0%; Score 22; DB 1; Length 2564;
 Best Local Similarity 100.0%; Pred. No. 8.6e-13;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 37 DERAIVQKKTFTKWNHSHLAV 58
 Db 57 DERAIVQKKTFTKWNHSHLAV 78

```

RESULT 41
ID Q71S07 HUMAN PRELIMINARY; PRT; 2564 AA.
AC Q71S07;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Non-erythrocytic beta-spectrin 4.
GN Name=SPNB4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
OC Homo.
OC NCBI_TaxID=9606;
OX NCBI_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
RA Cal T.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AF34063; AAQ14860.1; -; mRNA.
DR HSSP; Q01082; 1AA2.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR001589; Actinin actin bd.
DR InterPro; IPR001715; Calponin act_bd.
DR InterPro; IPR001849; PH.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001605; Spectrin_PH.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00435; Spectrin; 17.
DR PRINTS; PR00683; SPECTRINPH.
DR SMART; SM00033; CH; 2.
DR SMART; SM00233; PH; 1.
DR SMART; SM00150; SPEC; 16.
DR PROSITE; PS00019; ACTININ_1; 1.

```

```

DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
DR PROSITE; PS50003; PH DOMAIN; 1.
KW Actin capping; Actin-binding; Cytoskeleton.
SQ SEQUENCE 2564 AA; 289061 MW; CB1CABC2FF977776 CRC64;

```

Query Match 10.0%; Score 22; DB 2; Length 2564;
 Best Local Similarity 100.0%; Pred. No. 8.6e-13;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 37 DERAIVQKKTFTKWNHSHLAV 58
 Db 57 DERAIVQKKTFTKWNHSHLAV 78

```

RESULT 42
ID Q6NNX2 DROME PRELIMINARY; PRT; 1410 AA.
AC Q6NNX2;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE At2441p (Fragment).
GN Name=Beta-Spec;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN NUCLEOTIDE SEQUENCE.
RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise B.,
RA George R., Gonzalez W., Guarin H., Krommiller B., Li P., Liso G.,
RA Miranda A., Murgall C.J., Nunoo J., Pacle J., Patagas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BT011160; AAR82828.1; -; mRNA.
DR HSSP; Q01082; 1AA2.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR001589; Actbind actinin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00435; Spectrin; 10.
DR SMART; SM00033; CH; 2.
DR SMART; SM00150; SPEC; 10.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
FT NON TER 1410 1410
SQ SEQUENCE 1410 AA; 164827 MW; 3346E537C78AA48B CRC64;

```

Query Match 8.6%; Score 19; DB 2; Length 1410;
 Best Local Similarity 100.0%; Pred. No. 6.1e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 75 LKLEVLVSGERLPKPTKG 93
 Db 84 LKLEVLVSGERLPKPTKG 102

```

RESULT 43
ID SPTCB DROME STANDARD; PRT; 2291 AA.
AC Q00963; Q9VX30;
DT 01-APR-1993 (Rel. 25, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Spectrin beta chain.
GN Name=Beta-Spec; Synonyms=Spec-b; ORFNames=CG5870;
OS Drosophila melanogaster (Fruit fly).

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_Taxid=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA].
 RX MEDLINE=92335263; PubMed=1631106;
 RA Byers T.J., Brandon B., Lue R., Winograd E., Branton D.;
 RT "The complete sequence of Drosophila beta-spectrin reveals supra-
 RT motifs comprising eight 106-residue segments."
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6187-6191(1992).
 [2]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRIN-Berkely;
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blake R.G., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., McKlos G.L.G.,
 RA Abril J.F., Abmayant A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale U., Bayraktaroglu U., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Horin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Jaitani B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkars R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 [3]
 RN GENOME REANNOTATION.
 RP MEDLINE=22426069; PubMed=12537572;
 RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu U., Berman B.P.,
 RA Batencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review."
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [4]
 RN CHARACTERIZATION.
 RP MEDLINE=86059242; PubMed=3660372; DOI=10.1083/jcb.105.5.2095;
 RA Dubreuil R., Byers T.J., Branton D., Goldstein L.S.B., Kiehart D.P.;
 RT "Drosophila spectrin. I. Characterization of the purified protein."
 RL J. Cell Biol. 105:2095-2102(1987).
 [5]

RP STRUCTURE BY NMR OF 2145-2262.
 RX MEDLINE=96164435; PubMed=8591029; DOI=10.1016/S0969-2126(01)00254-4;
 RA Zhang P., Talluri S., Deng H., Branton D., Wagner G.;
 RT "Solution structure of the pleckstrin homology domain of Drosophila
 RT beta-spectrin."
 RL Structure 3:1185-1195(1995).
 CC -1- FUNCTION: Spectrin is the major constituent of the cytoskeletal
 CC network underlying the erythrocyte plasma membrane. It associates
 CC with band 4.1 and actin to form the cytoskeletal superstructure of
 CC the erythrocyte plasma membrane. Interacts with calmodulin in a
 CC calcium-dependent manner.
 CC -1- SUBUNIT: Native spectrin molecule is a tetramer composed of two
 CC antiparallel heterodimers joined head to head so that each end of
 CC the native molecule includes the C-terminus of the alpha subunit
 CC and the N-terminus of the beta subunit.
 CC -1- INTERACTION:
 CC Q9YTA2:CGI0627; Nbxp=1; Inlact=EBI-75025, EBI-119881;
 CC -1- SIMILARITY: Belongs to the spectrin family.
 CC -1- SIMILARITY: Contains 2 CH (calponin-homology) domains.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -1- SIMILARITY: Contains 17 spectrin repeats.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; M92288; AAA28399.1; -; mRNA.
 CC EMBL; AE003506; AAF48751.1; -; Genomic DNA.
 CC PIR; A46147; A46147.
 CC PDB; 1DRO; NMR; @2145-2262.
 CC Inlact; Q00963; -;
 CC Eneemb; CG5870; Drosophila melanogaster.
 CC Flybase; FBgn003471; Beta-Spec.
 CC GO; GO:0016327; C:apicolateral plasma membrane; TAS.
 CC GO; GO:0045169; C:fusome; IDA.
 CC GO; GO:0045170; C:spectrosome; IDA.
 CC GO; GO:0008017; P:microtubule binding; IDA.
 CC GO; GO:0045478; P:fusome organization and biogenesis; TAS.
 CC GO; GO:0007274; P:neuromuscular synaptic transmission; IMP.
 CC GO; GO:0007009; P:plasma membrane organization and biogenesis; TAS.
 CC GO; GO:0030721; P:spectrosome organization and biogenesis; TAS.
 CC InterPro; IPR001589; Actinin actin bd.
 CC InterPro; IPR001715; Calponin act_bd.
 CC InterPro; IPR001849; PH.
 CC InterPro; IPR011993; PH type.
 CC InterPro; IPR002017; Spectrin.
 CC InterPro; IPR001605; Spectrin_PH.
 CC Pfam; PF00307; CH; 2.
 CC Pfam; PF00169; PH; 1.
 CC Pfam; PF00435; Spectrin; 17.
 CC PRINTS; PR00683; SPECTRINPH.
 CC SMART; SMO0033; CH; 2.
 CC SMART; SMO0233; PH; 1.
 CC SMART; SMO0150; SPC; 17.
 CC PROSITE; PS00019; ACTININ_1; 1.
 CC PROSITE; PS00021; CH; 2.
 CC PROSITE; PS00020; ACTININ_2; 1.
 CC PROSITE; PS50003; PH DOMAIN; 1.
 CC 3D-structure; Actin binding; Actin-binding; Calmodulin-binding;
 CC Cytoskeleton; Membrane; Repeat.
 CC DOMAIN 1 271 Actin-binding (By similarity).
 CC FT DOMAIN 50 154 CH 1.
 CC FT DOMAIN 169 271 CH 2.
 CC FT REPEAT 298 408 Spectrin 1.
 CC FT REPEAT 418 522 Spectrin 2.
 CC FT REPEAT 524 633 Spectrin 3.
 CC FT REPEAT 635 739 Spectrin 4.
 CC FT REPEAT 741 844 Spectrin 5.
 CC FT REPEAT 846 950 Spectrin 6.
 CC FT REPEAT 952 1057 Spectrin 7.
 CC FT REPEAT 1059 1167 Spectrin 8.

```
FT REPEAT 1169 1273 Spectrin 9.
FT REPEAT 1275 1378 Spectrin 10.
FT REPEAT 1380 1485 Spectrin 11.
FT REPEAT 1487 1591 Spectrin 12.
FT REPEAT 1593 1697 Spectrin 13.
FT REPEAT 1699 1804 Spectrin 14.
FT REPEAT 1806 1910 Spectrin 15.
FT REPEAT 1912 2016 Spectrin 16.
FT REPEAT 2018 2078 Spectrin 17.
FT DOMAIN 2147 2259 PH.
FT CONFLICT 2278 2278 D -> Y (in Ref. 1).
FT STRAND 2150 2157
FT TURN 2166 2167
FT STRAND 2173 2179
FT STRAND 2185 2187
FT HELIX 2192 2195
FT TURN 2197 2198
FT STRAND 2206 2206
FT STRAND 2215 2216
FT TURN 2224 2225
FT STRAND 2226 2230
FT STRAND 2236 2240
FT HELIX 2245 2258
FT TURN 2259 2259
SQ SEQUENCE 2291 AA; 265739 MW; 5CDPB0C548BHC39B CRC64;

Query Match 8.6%; Score 19; DB 1; Length 2291;
Best Local Similarity 100.0%; Pred. No. 9,7e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 75 LKLEVLGSGRLPPTKG 93
Db 84 LKLEVLGSGRLPPTKG 102

RESULT 44
O9TR82_PIG PRELIMINARY; PRT; 20 AA.
ID O9TR82_PIG
AC O9TR82_PIG
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE Spectrin beta SUBUNIT=28 kDa fragment (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suidae;
OC Sus.
OX NCBI_TaxID=9623;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=92209538; PubMed=1555607;
RA Firepiller T., Derancourt J., Pradel L.A.;
RT "Actin and neurofilament binding domain of brain spectrin beta
RT subunit."
RL Eur. J. Biochem. 205:85-91(1992).
DR HSSP; Q15149; IMB8.
DR GO; GO:0003779; F:actin binding; IEA.
DR InterPro; IPR001589; Actbind actnm.
DR PROSITE; PS00019; ACTININ_1; 1.
FT NON_TER 1 1
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2322 MW; F23D09A9EAD01F1E CRC64;

Query Match 7.7%; Score 17; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1,2e-09;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 37 DERBAVOKKTKTKWNS 53
Db 4 DERBAVOKKTKTKWNS 20
```

```
RESULT 45
O28297_CANFA PRELIMINARY; PRT; 707 AA.
ID O28297_CANFA
AC O28297;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Beta-spectrin (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC Canis.
OX NCBI_TaxID=9615;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Trachea;
RX MEDLINE=93176817; PubMed=8439565; DOI=10.1016/0167-4781(93)90299-S;
RA Tan S., Shankar V., Gilmore M.S., Sachdev G.P.;
RT "Nucleotide sequence of a cDNA for canine beta-spectrin reveals high
RT evolutionary conservation."
RL Biochim. Biophys. Acta 1172:217-219(1993).
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; L02897; AAA30879.1; -; mRNA.
DR PIR; S29854; S29854.
DR HSSP; Q62261; 1BTN.
DR SMR; Q28297; 543-648.
DR Ensembl; ENSGAF00000002744; Canis familiaris.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011993; PH type.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001605; Spectrin_PH.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00435; Spectrin; 5.
DR PRINTS; PR00683; SPECTRINPH.
DR SMART; SM00233; PH; 1.
DR SMART; SM00150; SPEC; 4.
DR PROSITE; PS50003; PH_DOMAIN; 1.
FT NON_TER 1 1
FT NON_TER 707 707
SQ SEQUENCE 707 AA; 8118 MW; DID6C6CD2108D459 CRC64;

Query Match 7.3%; Score 16; DB 2; Length 707;
Best Local Similarity 100.0%; Pred. No. 3,9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 141 RROQEEERKRRPSP 156
Db 431 RROQEEERKRRPSP 446

RESULT 46
O7PSH4_ANOGA PRELIMINARY; PRT; 2296 AA.
ID O7PSH4_ANOGA
AC O7PSH4;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ENSANGP00000012507.
GN ENSANGP00000012507.
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL "Anopheles gambiae re-annotation."
RN Submitted (APR-2002) to the EMBL/Genbank/DBJ databases.
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
```

RG The Anopheles gambiae Sequence Committee;
 RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAAB01008823; EAA05634.2; -; Genomic_DNA.
 DR HSSP; Q01082; 1BKR.
 DR GO; GO:0003779; F:actin binding; IEA.
 DR InterPro; IPR001589; Actinin actin bd.
 DR InterPro; IPR001715; Calponin_act_bd.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR011993; PH_type.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR001605; Spectrin_PH.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00435; Spectrin_17.
 DR PRINTS; PR00683; SPECTRINPH.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 KW Actin capping; Actin-binding; Cytoskeleton.
 SQ SEQUENCE 2296 AA; 266023 MW; 1B1974E0F7F4E0 CRC64;
 Query Match 7.3%; Score 16; DB 2; Length 2296;
 Best Local Similarity 100.0%; Pred.No. 1,2e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 MRHCLENVKXALQPL 110
 Db 104 MRHCLENVKXALQPL 119
 RESULT 47
 Q4RFW3_TETNG PRELIMINARY; PRT; 129 AA.
 AC Q4RFW3;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Chromosome 16 SCAP15113, whole genome shotgun sequence.
 OS ORFNames=GSTENG00035151001;
 GN Tetradon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Butelosteii; Neocentropomidae;
 OC Acanthomorphae; Acanthopterygii; Percomorphae; Tetradontiformes;
 OC Tetradontidae; Tetradontidae; Tetradon.
 CC NCBI_TaxID=99883;
 RX NCBI_TaxID=99883;
 RP NUCLEOTIDE SEQUENCE.
 RA Jallion O., Arny J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Maucelli E., Bonnaud U., Fischer C., Ozout-Costaz C., Bernot A.,
 RA Nicud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Daetliwa C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Antouard V., Jubin C., Castelli V., Katinka M., Vachelle B.,
 RA Bismont C., Skaili Z., Catolico L., Poulin J., De Bernardis V.,
 RA Cruaud C., Duprat S., Broctier P., Coutanceau J.P., Guzy J.,
 RA Patra G., Lardier G., Chappe C., McKernan K.J., McEwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Lauder V., Schachter V., Queller F., Saurin W., Scarpelli C.,
 RA Winkler P., Lander B.S., Weissbach J., Roest Crolius H.;
 RA "Genome duplication in the teleost fish Tetradon nigroviridis reveals
 RT the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (Feb-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL; CAAB01015113; CAG12719.1; -; Genomic DNA.
 SQ SEQUENCE 129 AA; 14599 MW; CSB34F14E04E7E1C CRC64;
 Query Match 6.8%; Score 15; DB 2; Length 129;
 Best Local Similarity 100.0%; Pred.No. 8,1e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 41 AVOKKTFKWNVSHL 55
 Db 53 AVOKKTFKWNVSHL 67
 RESULT 48
 Q9U9J7_CAEEL PRELIMINARY; PRT; 2257 AA.
 ID Q9U9J7_CAEEL
 AC Q9U9J7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Beta-G spectrin (Beta-spectrin) (uncoordinated protein 70, isoform
 DE a).
 CN Name=bgs-1; Synonyms=unc-70; ORFNames=K11C4.3;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 CC NCBI_TaxID=6239;
 RX NCBI_TaxID=6239;
 RP NUCLEOTIDE SEQUENCE.
 RA Moorthy S., Chen L., Bennett V.,
 RT "Caenorhabditis elegans beta-g spectrin is dispensable for
 RT establishment of epithelial polarity, but essential for muscular and
 RT neuronal function.";
 RL J. Cell. Biochem. 149:915-930(2000).
 [2]
 RN NUCLEOTIDE SEQUENCE.
 RP MEDLINE=20273916; PubMed=10811832; DOI=10.1083/jcb.149.4.931;
 RX Hammariund M., Davis W.S., Jorgensen E.M.;
 RT "Mutations in beta-spectrin disrupt axon outgrowth and sarcomere
 RT structure.";
 RL J. Cell Biol. 149:931-942(2000).
 [3]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=Bristol N2;
 RC MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 CC -!- SIMILARITY: Contains 1 PH domain.
 DR EMBL; AF16170; AAD49859.1; -; mRNA.
 DR EMBL; AF261891; AAF72996.1; -; mRNA.
 DR EMBL; U64854; AAK77611.1; -; Genomic DNA.
 DR HSSP; Q01082; 1BKR.
 DR Ensemble; K11C4.3; Caenorhabditis elegans.
 DR WormBase; WBGene0006803; K11C4.3.
 DR GO; GO:0003779; F:actin binding; IEA.
 DR GO; GO:0003779; F:actin binding; IEA.
 DR InterPro; IPR001589; Actinin actin bd.
 DR InterPro; IPR001715; Calponin_act_bd.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR011993; PH_type.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR001605; Spectrin_PH.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00435; Spectrin_17.
 DR PRINTS; PR00683; SPECTRINPH.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00150; SPC; 17.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.

DR PROSITE; PS50003; PH DOMAIN; 1.
 KW Actin capping; Actin-binding; Complete proteome; Cytoskeleton.
 SQ SEQUENCE 2257 AA; 262228 MW; E1393B6D0914D803 CRC64;

Query Match 6.4%; Score 14; DB 2; Length 2257;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTTKVNSHL 55
 |||||
 DB 41 VOKKFTTKVNSHL 54

RESULT 49
 Q60WP8 CAEBR PRELIMINARY; PRT; 2299 AA.

ID Q60WP8 CAEBR PRELIMINARY; PRT; 2299 AA.
 AC Q60WP8;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Hypothetical protein CBG19038 (Fragment).
 GN Name=CBG19038;
 OS Caenorhabditis briggsae.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 CX NCBI_Taxid=6238;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RG The C. briggsae Sequencing Consortium;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 PH domain.
 DR EMBL; CAC01000098; CAB1967.1; -; Genomic_DNA.
 DR GO; GO:0003779; F:actin binding; IEA.
 DR InterPro; IPR001589; Actinin_actin_bd.
 DR InterPro; IPR001715; Calponin_act_bd.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR001605; Spectrin_PH.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00435; Spectrin; 17.
 DR PRINTS; PR00683; SPECTRINPH.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00150; SPEC; 17.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS50003; PH DOMAIN; 1.
 KW Actin capping; Actin-binding; Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 2299 AA; 267209 MW; 356A42856E3A4462 CRC64;

Query Match 6.4%; Score 14; DB 2; Length 2299;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTTKVNSHL 55
 |||||
 DB 38 VOKKFTTKVNSHL 51

RESULT 50
 Q95ZL8 CAEBL PRELIMINARY; PRT; 2302 AA.

ID Q95ZL8 CAEBL PRELIMINARY; PRT; 2302 AA.
 AC Q95ZL8;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Uncoordinated protein 70, isoform b.
 GN Name=unc-70; ORFNames=K11C4.3;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.
 CX NCBI_Taxid=6239;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG The C. elegans sequencing consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology";
 RL Science 282:2012-2018(1998).

CC -1- SIMILARITY: Contains 1 PH domain.
 DR EMBL; U64854; AAK7612.2; -; Genomic_DNA.
 DR HSSP; Q01082; 1BKR.
 DR Ensembl; K11C4.3; Caenorhabditis elegans.
 DR WormBase; WBGene00006803; K11C4.3.
 DR WormPep; K11C4.3b; CE30159.
 DR GO; GO:0003779; F:actin binding; IEA.
 DR InterPro; IPR001589; Actinin_actin_bd.
 DR InterPro; IPR001715; Calponin_act_bd.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR011993; PH_type.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR001605; Spectrin_PH.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00435; Spectrin; 17.
 DR PRINTS; PR00683; SPECTRINPH.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00150; SPEC; 17.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.
 DR PROSITE; PS50003; PH DOMAIN; 1.
 KW Actin capping; Actin-binding; Complete proteome; Cytoskeleton.
 FT NON_TER 1
 SQ SEQUENCE 2302 AA; 267285 MW; 17E29A016B72E26B CRC64;

Query Match 6.4%; Score 14; DB 2; Length 2302;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTTKVNSHL 55
 |||||
 DB 41 VOKKFTTKVNSHL 54

Search completed: March 4, 2006, 07:10:11
 Job time : 285 sec

This Page Blank (uspio)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2006, 07:11:46 ; Search time 20 Seconds
(without alignments)
220.005 Million cell updates/sec

Title: US-10-695-994A-7

Perfect score: 220

Sequence: 1 MEIQRTRSSIGSLPAYTQ.....QNYKNFNSRRRTASDHWSGM 220

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 135346 seqs, 20000420 residues

Word size: 6

Total number of hits satisfying chosen parameters: 183

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Published Applications_AA_New:*
1: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	11	5.0	1452	6 US-10-995-561-778	Sequence 778, App
2	11	5.0	3803	6 US-10-995-561-773	Sequence 772, App
3	11	5.0	3960	6 US-10-995-561-771	Sequence 771, App
4	11	5.0	5406	6 US-10-995-561-774	Sequence 774, App
5	11	5.0	5415	6 US-10-995-561-779	Sequence 779, App
6	11	5.0	5464	6 US-10-995-561-775	Sequence 775, App
7	9	4.1	15	7 US-11-106-932-113	Sequence 113, App
8	9	4.1	30	7 US-11-106-932-33	Sequence 33, Appl
9	8	3.6	919	6 US-10-821-224-951	Sequence 951, App
10	8	3.6	1089	7 US-11-087-099-7653	Sequence 7653, App
11	8	3.6	1609	7 US-11-087-099-4181	Sequence 4181, App
12	7	3.2	186	7 US-11-072-512-3726	Sequence 3726, App
13	7	3.2	306	6 US-10-995-561-702	Sequence 702, App
14	7	3.2	312	6 US-10-995-561-700	Sequence 700, App
15	7	3.2	361	6 US-10-995-561-701	Sequence 701, App
16	7	3.2	364	6 US-10-995-561-699	Sequence 699, App
17	7	3.2	549	7 US-11-072-512-3479	Sequence 3479, App
18	7	3.2	595	7 US-11-087-099-2273	Sequence 2273, App
19	7	3.2	673	7 US-11-072-512-2425	Sequence 2425, App
20	7	3.2	1099	6 US-10-873-528-111	Sequence 111, Appl
21	7	3.2	4374	7 US-11-128-572-2	Sequence 2, Appl
22	7	3.2	5335	6 US-10-995-561-777	Sequence 777, App
23	6	2.7	8	7 US-11-045-024-2635	Sequence 2635, App
24	6	2.7	8	7 US-11-045-024-5011	Sequence 5011, App
25	6	2.7	9	7 US-11-045-024-2649	Sequence 2649, App

26	6	2.7	9	7 US-11-045-024-3576	Sequence 3576, App
27	6	2.7	9	7 US-11-045-024-5023	Sequence 5023, App
28	6	2.7	9	7 US-11-045-024-6897	Sequence 6897, App
29	6	2.7	9	7 US-11-045-024-10720	Sequence 10720, App
30	6	2.7	9	7 US-11-045-024-12496	Sequence 12496, App
31	6	2.7	9	7 US-11-045-024-14001	Sequence 14001, App
32	6	2.7	10	7 US-11-045-024-3582	Sequence 3582, App
33	6	2.7	10	7 US-11-045-024-3582	Sequence 3582, App
34	6	2.7	10	7 US-11-045-024-10716	Sequence 10716, App
35	6	2.7	10	7 US-11-045-024-12492	Sequence 12492, App
36	6	2.7	10	7 US-11-045-024-12853	Sequence 12853, App
37	6	2.7	15	7 US-11-045-024-13345	Sequence 13345, App
38	6	2.7	15	7 US-11-045-024-13348	Sequence 13348, App
39	6	2.7	15	7 US-11-045-024-13515	Sequence 13515, App
40	6	2.7	15	7 US-11-045-024-13516	Sequence 13516, App
41	6	2.7	35	7 US-11-096-725-28	Sequence 28, Appl
42	6	2.7	100	7 US-11-107-096-78	Sequence 78, Appl
43	6	2.7	109	7 US-11-053-076-59	Sequence 59, Appl
44	6	2.7	113	6 US-10-485-788A-724	Sequence 724, Appl
45	6	2.7	113	7 US-11-053-076-94	Sequence 94, Appl
46	6	2.7	126	7 US-11-055-822-146	Sequence 146, App
47	6	2.7	134	7 US-11-052-554A-88	Sequence 88, Appl
48	6	2.7	145	7 US-11-087-099-7788	Sequence 7788, App
49	6	2.7	163	6 US-10-510-386-194	Sequence 194, App
50	6	2.7	165	7 US-11-087-099-5484	Sequence 5484, App
51	6	2.7	184	7 US-11-084-591-5	Sequence 5, Appl
52	6	2.7	184	7 US-11-087-099-5689	Sequence 5689, App
53	6	2.7	185	6 US-10-793-626-2102	Sequence 2102, App
54	6	2.7	191	7 US-11-176-830-875	Sequence 875, App
55	6	2.7	215	7 US-11-089-551A-15	Sequence 15, Appl
56	6	2.7	226	6 US-10-511-538-194	Sequence 194, App
57	6	2.7	237	7 US-11-083-055-2	Sequence 2, Appl
58	6	2.7	246	7 US-11-098-686-11081	Sequence 11081, App
59	6	2.7	247	7 US-11-098-686-10406	Sequence 10406, App
60	6	2.7	251	7 US-11-098-686-10800	Sequence 10800, App
61	6	2.7	264	6 US-10-467-657-1332	Sequence 1332, App
62	6	2.7	264	7 US-11-087-099-4412	Sequence 4412, App
63	6	2.7	271	7 US-11-089-551A-17	Sequence 17, Appl
64	6	2.7	281	7 US-11-087-099-6709	Sequence 6709, App
65	6	2.7	281	7 US-11-087-099-7386	Sequence 7386, App
66	6	2.7	284	6 US-10-873-528-40	Sequence 40, Appl
67	6	2.7	286	7 US-11-087-099-4420	Sequence 4420, App
68	6	2.7	286	7 US-11-087-099-4942	Sequence 4942, App
69	6	2.7	286	7 US-11-087-099-6943	Sequence 6943, App
70	6	2.7	286	7 US-11-087-099-7513	Sequence 7513, App
71	6	2.7	287	7 US-11-087-099-8943	Sequence 8943, App
72	6	2.7	287	7 US-11-087-099-9180	Sequence 9180, App
73	6	2.7	287	7 US-11-087-099-8561	Sequence 8561, App
74	6	2.7	287	7 US-11-087-099-11384	Sequence 11384, App
75	6	2.7	296	7 US-11-074-176-22	Sequence 22, Appl
76	6	2.7	310	7 US-11-087-099-11105	Sequence 11105, App
77	6	2.7	322	7 US-11-073-605-4	Sequence 4, Appl
78	6	2.7	322	7 US-11-140-284-6	Sequence 6, Appl
79	6	2.7	324	7 US-11-087-099-10611	Sequence 10611, App
80	6	2.7	331	7 US-11-087-099-9511	Sequence 9511, App
81	6	2.7	333	7 US-11-087-099-490	Sequence 490, App
82	6	2.7	333	7 US-11-087-099-6265	Sequence 6265, App
83	6	2.7	336	7 US-11-087-099-8846	Sequence 8846, App
84	6	2.7	354	7 US-11-087-099-6791	Sequence 6791, App
85	6	2.7	355	7 US-11-087-099-994	Sequence 994, App
86	6	2.7	366	6 US-11-098-686-10689	Sequence 10689, App
87	6	2.7	366	6 US-10-131-826A-186	Sequence 186, App
88	6	2.7	364	6 US-10-973-115B-186	Sequence 186, App
89	6	2.7	370	7 US-11-073-605-2	Sequence 2, Appl
90	6	2.7	370	7 US-11-073-605-14	Sequence 14, Appl
91	6	2.7	370	7 US-11-140-284-8	Sequence 8, Appl
92	6	2.7	376	6 US-10-517-939-102	Sequence 102, Appl
93	6	2.7	379	7 US-11-087-099-872	Sequence 872, App
94	6	2.7	387	7 US-11-087-099-11673	Sequence 11673, App
95	6	2.7	391	7 US-11-087-099-10130	Sequence 10130, App
96	6	2.7	409	7 US-11-087-099-2414	Sequence 2414, App
97	6	2.7	411	7 US-11-092-168-8	Sequence 8, Appl

```

99 2.7 421 7 US-11-072-512-3536
100 2.7 429 7 US-11-053-185-2
101 2.7 432 6 US-10-467-657-8783
102 2.7 445 6 US-10-821-234-1506
103 2.7 451 7 US-11-087-099-4574
104 2.7 451 7 US-11-186-284-79
105 2.7 452 7 US-11-087-099-2368
106 2.7 452 7 US-11-087-099-8208
107 2.7 455 5 US-09-978-360A-523
108 2.7 459 6 US-11-087-099-8118
109 2.7 465 6 US-10-793-628-2828
110 2.7 468 6 US-10-957-589-28
111 2.7 468 6 US-11-097-589-26
112 2.7 470 6 US-10-878-556A-101
113 2.7 471 7 US-11-098-686-11324
114 2.7 478 7 US-11-098-686-10246
115 2.7 479 6 US-10-821-234-871
116 2.7 480 7 US-11-087-099-8522
117 2.7 481 6 US-10-508-263-28
118 2.7 485 7 US-11-072-512-3419
119 2.7 487 7 US-11-072-512-2868
120 2.7 498 7 US-11-072-512-3191
121 2.7 499 7 US-11-087-099-1159
122 2.7 499 7 US-11-087-099-5941
123 2.7 516 6 US-10-508-263-26
124 2.7 518 7 US-11-072-512-3059
125 2.7 539 6 US-10-467-657-6792
126 2.7 562 6 US-10-508-263-24
127 2.7 567 6 US-10-330-773-556
128 2.7 582 6 US-10-330-773-558
129 2.7 612 6 US-10-467-657-764
130 2.7 620 6 US-10-330-773-75
131 2.7 629 7 US-11-098-686-11359
132 2.7 643 7 US-11-087-099-1991
133 2.7 662 7 US-11-072-175-184
134 2.7 664 6 US-10-878-556A-44
135 2.7 685 7 US-11-089-551A-2
136 2.7 715 7 US-11-087-099-2922
137 2.7 722 6 US-10-858-730-77
138 2.7 792 7 US-11-087-099-1215
139 2.7 805 6 US-10-485-517-198
140 2.7 853 6 US-10-420-192-6
141 2.7 908 6 US-10-821-234-1267
142 2.7 909 7 US-11-077-619-8
143 2.7 914 7 US-11-072-512-2923
144 2.7 923 7 US-11-087-099-9592
145 2.7 936 7 US-11-072-512-2621
146 2.7 1045 7 US-11-072-512-2830
147 2.7 1047 7 US-11-072-512-2408
148 2.7 1085 7 US-11-087-099-11646
149 2.7 1087 7 US-11-117-169-10
150 2.7 1106 7 US-11-076-427A-16
151 2.7 1106 7 US-11-075-047A-119
152 2.7 1123 7 US-11-019-711-71
153 2.7 1147 7 US-11-087-099-11762
154 2.7 1148 7 US-11-110-082-29
155 2.7 1149 7 US-11-110-082-30
156 2.7 1327 7 US-11-019-711-70
157 2.7 1342 7 US-11-115-639-2
158 2.7 1342 7 US-11-115-639-3
159 2.7 1342 7 US-11-115-639-4
160 2.7 1342 7 US-11-115-639-5
161 2.7 1342 7 US-11-115-639-6
162 2.7 1347 7 US-11-087-099-5370
163 2.7 1360 7 US-11-241-056-14
164 2.7 1381 6 US-10-467-657-178
165 2.7 1381 6 US-10-467-657-3726
166 2.7 1562 7 US-11-052-554A-211
167 2.7 1565 6 US-10-467-657-2704
168 2.7 1681 7 US-11-019-711-20
169 2.7 1697 7 US-11-019-711-68
170 2.7 1723 7 US-11-019-711-18
171 2.7 1897 6 US-10-821-234-1635

```

```

Sequence 3536, Ap
Sequence 2, Appl
Sequence 8783, Ap
Sequence 1606, Ap
Sequence 4574, Ap
Sequence 79, Appl
Sequence 2368, Ap
Sequence 8208, Ap
Sequence 523, Ap
Sequence 8118, Ap
Sequence 2928, Ap
Sequence 28, Appl
Sequence 26, Appl
Sequence 101, App
Sequence 11324, A
Sequence 10246, A
Sequence 871, App
Sequence 8522, Ap
Sequence 28, Appl
Sequence 3419, Ap
Sequence 2868, Ap
Sequence 3191, Ap
Sequence 1159, Ap
Sequence 5941, Ap
Sequence 26, Appl
Sequence 3059, Ap
Sequence 6792, Ap
Sequence 24, Appl
Sequence 556, App
Sequence 558, App
Sequence 764, App
Sequence 75, Appl
Sequence 11359, A
Sequence 1991, Ap
Sequence 184, App
Sequence 44, Appl
Sequence 2, Appl
Sequence 2922, Ap
Sequence 77, Appl
Sequence 1215, Ap
Sequence 198, App
Sequence 6, Appl
Sequence 1267, Ap
Sequence 8, Appl
Sequence 2923, Ap
Sequence 2621, Ap
Sequence 2830, Ap
Sequence 2408, Ap
Sequence 11646, A
Sequence 10, Appl
Sequence 16, Appl
Sequence 119, App
Sequence 71, Appl
Sequence 11762, A
Sequence 29, Appl
Sequence 30, Appl
Sequence 2, Appl
Sequence 3, Appl
Sequence 5, Appl
Sequence 6, Appl
Sequence 5370, Ap
Sequence 14, Appl
Sequence 178, Appl
Sequence 3726, Ap
Sequence 211, App
Sequence 2704, Ap
Sequence 20, Appl
Sequence 68, Appl
Sequence 18, Appl
Sequence 1635, Ap

```

```

172 2.7 1907 7 US-11-000-463-250 Sequence 250, App
173 2.7 1976 7 US-11-069-834-52 Sequence 52, Appl
174 2.7 1976 7 US-11-069-834-54 Sequence 54, Appl
175 2.7 2117 7 US-11-087-099-9594 Sequence 9594, Ap
176 2.7 2515 7 US-11-113-424-53 Sequence 53, Appl
177 2.7 2640 7 US-11-087-099-11966 Sequence 11966, A
178 2.7 2647 6 US-10-821-234-1303 Sequence 1303, Ap
179 2.7 2657 6 US-10-821-234-1262 Sequence 1262, Ap
180 2.7 3674 7 US-11-000-463-454 Sequence 454, App
181 2.7 4384 6 US-10-821-234-1120 Sequence 1120, Ap
182 2.7 5291 7 US-11-052-554A-281 Sequence 281, App
183 2.7 8746 7 US-11-098-686-10232 Sequence 10232, A

```

ALIGNMENTS

```

RESULT 1
US-10-995-561-778
; Sequence 778, Application US/10995561
; Publication No. US20050272054A1
GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 778
; LENGTH: 1452
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-995-561-778

Query Match
Best Local Similarity 5.0%; Score 11; DB 6; Length 1452;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKTFTKWN 52
|||||
Db 79 VOKKTFTKWN 89

RESULT 2
US-10-995-561-773
; Sequence 773, Application US/10995561
; Publication No. US20050272054A1
GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 773
; LENGTH: 3803
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-995-561-773

Query Match
Best Local Similarity 5.0%; Score 11; DB 6; Length 3803;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKTFTKWN 52
|||||
Db 79 VOKKTFTKWN 89

```

```
RESULT 3
US-10-995-561-771
; Sequence 771, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 771
; LENGTH: 3960
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-771

Query Match
Best Local Similarity 5.0%; Score 11; DB 6; Length 3960;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VQKKTFTKWN 52
DB 79 VQKKTFTKWN 89

RESULT 4
US-10-995-561-774
; Sequence 774, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 774
; LENGTH: 5406
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-774

Query Match
Best Local Similarity 5.0%; Score 11; DB 6; Length 5406;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VQKKTFTKWN 52
DB 79 VQKKTFTKWN 89

RESULT 5
US-10-995-561-779
; Sequence 779, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 779
; LENGTH: 5415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-779

Query Match
Best Local Similarity 5.0%; Score 11; DB 6; Length 5415;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VQKKTFTKWN 52
DB 79 VQKKTFTKWN 89

RESULT 6
US-10-995-561-775
; Sequence 775, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: C1001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 775
; LENGTH: 5464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-775

Query Match
Best Local Similarity 5.0%; Score 11; DB 6; Length 5464;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VQKKTFTKWN 52
DB 79 VQKKTFTKWN 89

RESULT 7
US-11-106-932-113
; Sequence 113, Application US/11106932
; Publication No. US20050260697A1
; GENERAL INFORMATION:
; APPLICANT: WANG, KA-WANG KEVIN
; APPLICANT: HAYES, RONALD
; APPLICANT: LIU, MING CHEN
; APPLICANT: OLI, MONIKA
; TITLE OF INVENTION: PROTEOLYTIC MARKERS AS DIAGNOSTIC BIOMARKERS FOR CANCER, ORGAN
; TITLE OF INVENTION: INJURY AND MUSCLE REHABILITATION/EXERCISE OVERTRAINING
; FILE REFERENCE: 5853-549-1
; CURRENT APPLICATION NUMBER: US/11/106,932
; CURRENT FILING DATE: 2005-04-15
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 113
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-106-932-113

Query Match
Best Local Similarity 4.1%; Score 9; DB 7; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0086;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 LPAEQSPR 192
|||||
```

Db 1 LPAEQGSPR 9

RESULT 8

US-11-106-932-33

Sequence 33, Application US/11106932

Publication No. US20050260697A1

GENERAL INFORMATION:

APPLICANT: WANG, KA-WANG KEVIN

APPLICANT: HAYES, RONALD

APPLICANT: LIU, MING CHEN

APPLICANT: OLI, MONIKA

TITLE OF INVENTION: PROTEOLYTIC MARKERS AS DIAGNOSTIC BIOMARKERS FOR CANCER, ORGAN

FILE REFERENCE: 5853-549-1

CURRENT APPLICATION NUMBER: US/11/106,932

CURRENT FILING DATE: 2005-04-15

NUMBER OF SEQ ID NOS: 147

SOFTWARE: PatentIn version 3.2

SEQ ID NO 33

LENGTH: 30

TYPE: PRT

ORGANISM: Homo sapiens

US-11-106-932-33

Query Match

Best Local Similarity 4.1%; Score 9; DB 7; Length 30;

Best Local Similarity 100.0%; Pred. No. 0.016;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 LPAEQGSPR 192

Db 1 LPAEQGSPR 9

RESULT 9

US-10-821-234-951

Sequence 951, Application US/10821234

Publication No. US20050255114A1

GENERAL INFORMATION:

APPLICANT: Labat, Ivan

APPLICANT: Stache-Crain, Birgit

APPLICANT: Andarmah, Susan

APPLICANT: Tang, Y. Tom

TITLE OF INVENTION: Methode for Diagnosis and Treatment of Preeclampsia

FILE REFERENCE: 821A

CURRENT APPLICATION NUMBER: US/10/821,234

CURRENT FILING DATE: 2004-04-07

PRIOR APPLICATION NUMBER: US 60/462,047

PRIOR FILING DATE: 2003-04-07

NUMBER OF SEQ ID NOS: 1704

SOFTWARE: pc_seq_genes Version 1.0

SEQ ID NO 951

LENGTH: 919

TYPE: PRT

ORGANISM: Homo sapiens

US-10-821-234-951

Query Match

Best Local Similarity 3.6%; Score 8; DB 6; Length 919;

Best Local Similarity 100.0%; Pred. No. 4.3;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 SGERLPKP 90

Db 96 SGERLPKP 103

RESULT 10

US-11-087-099-7653

Sequence 7653, Application US/11087099

Publication No. US20060041961A1

GENERAL INFORMATION:

APPLICANT: Abad, Mark S. et al.

TITLE OF INVENTION: Genes and Uses for Plant Improvement

FILE REFERENCE: 38-21(53450)B EP

CURRENT APPLICATION NUMBER: US/11/087,099

CURRENT FILING DATE: 2005-03-22

NUMBER OF SEQ ID NOS: 12464

SEQ ID NO 7653

LENGTH: 1089

TYPE: PRT

ORGANISM: Streptomyces coelicolor A3(2)

US-11-087-099-7653

Query Match

Best Local Similarity 3.6%; Score 8; DB 7; Length 1089;

Best Local Similarity 100.0%; Pred. No. 5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 EEBERKR 152

Db 761 EEBERKR 768

RESULT 11

US-11-087-099-4181

Sequence 4181, Application US/11087099

Publication No. US20060041961A1

GENERAL INFORMATION:

APPLICANT: Abad, Mark S. et al.

TITLE OF INVENTION: Genes and Uses for Plant Improvement

FILE REFERENCE: 38-21(53450)B EP

CURRENT APPLICATION NUMBER: US/11/087,099

CURRENT FILING DATE: 2005-03-22

NUMBER OF SEQ ID NOS: 12464

SEQ ID NO 4181

LENGTH: 1609

TYPE: PRT

ORGANISM: Neurospora crassa

US-11-087-099-4181

Query Match

Best Local Similarity 3.6%; Score 8; DB 7; Length 1609;

Best Local Similarity 100.0%; Pred. No. 7.2;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 OEBERKR 151

Db 1191 OEBERKR 1198

RESULT 12

US-11-072-512-3726

Sequence 3726, Application US/11072512

Publication No. US20060029945A1

GENERAL INFORMATION:

APPLICANT: ISOGAI, TAKAO

APPLICANT: SUGIYAMA, TOMOYASU

APPLICANT: OTSUKI, TENSUJI

APPLICANT: WAKAMATSU, AI

APPLICANT: SATO, HIROYUKI

APPLICANT: ISHII, SHIZUKO

APPLICANT: YAMAMOTO, JUN-ICHI

APPLICANT: ISONO, YUUKO

APPLICANT: HIO, YURI

APPLICANT: OTSUKA, KAORU

APPLICANT: NAGAI, KEIICHI

APPLICANT: IRIE, RYOTARO

APPLICANT: TAMECHIKA, ICHIRO

APPLICANT: SEKI, NAOHIKO

APPLICANT: YOSHIKAWA, TSUTOMU

APPLICANT: OTSUKA, MOTOTYUKI

APPLICANT: NAGAHARI, KENJI

APPLICANT: MASUHO, YASUHIKO

TITLE OF INVENTION: Novel full length cDNA

FILE REFERENCE: 08435-0191

CURRENT APPLICATION NUMBER: US/11/072,512

CURRENT FILING DATE: 2005-03-07

PRIOR APPLICATION NUMBER: US 60/350,978

PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3726
LENGTH: 186
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-3726

Query Match 3.2%; Score 7; DB 7; Length 186;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 152 RPPSPDP 158
Db 103 RPPSPDP 109

RESULT 13

US-10-995-561-702
Sequence 702, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 702
LENGTH: 306
TYPE: PRT
ORGANISM: Homo sapiens
US-10-995-561-702

Query Match 3.2%; Score 7; DB 6; Length 306;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 110 LKEQRVH 116
Db 144 LKEQRVH 150

RESULT 14

US-10-995-561-700
Sequence 700, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 700
LENGTH: 312
TYPE: PRT
ORGANISM: Homo sapiens
US-10-995-561-700

Query Match 3.2%; Score 7; DB 6; Length 312;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 110 LKEQRVH 116
Db 199 LKEQRVH 205

RESULT 15

US-10-995-561-701
Sequence 701, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 701
LENGTH: 361
TYPE: PRT
ORGANISM: Homo sapiens
US-10-995-561-701

Query Match 3.2%; Score 7; DB 6; Length 361;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 110 LKEQRVH 116
Db 199 LKEQRVH 205

RESULT 16

US-10-995-561-699
Sequence 699, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 699
LENGTH: 364
TYPE: PRT
ORGANISM: Homo sapiens
US-10-995-561-699

Query Match 3.2%; Score 7; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 110 LKEQRVH 116
Db 199 LKEQRVH 205

RESULT 17

US-11-072-512-3479
Sequence 3479, Application US/11072512
Publication No. US2006002945A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI

APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHITAKA, TSUTOMU
APPLICANT: OTSUKA, MOTOKYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072,512
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3479
LENGTH: 549
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-3479

Query Match 3.2%; Score 7; DB 7; Length 549;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 EQRVHLE 118
DB 28 EQRVHLE 34

RESULT 18
US-11-087-099-2273
Sequence 2273, Application US/11087099
Publication No. US20060041961A1
GENERAL INFORMATION:
APPLICANT: Abad, Mark S. et al.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)B BP
CURRENT APPLICATION NUMBER: US/11/087,099
CURRENT FILING DATE: 2005-03-22
NUMBER OF SEQ ID NOS: 12464
SEQ ID NO 2273
LENGTH: 595
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-11-087-099-2273

Query Match 3.2%; Score 7; DB 7; Length 595;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 EYLSGER 86
DB 252 EYLSGER 258

RESULT 19
US-11-072-512-2425
Sequence 2425, Application US/11072512
Publication No. US20060029945A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI

APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHITAKA, TSUTOMU
APPLICANT: OTSUKA, MOTOKYUKI
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072,512
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2425
LENGTH: 673
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-2425

Query Match 3.2%; Score 7; DB 7; Length 673;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 LKLELV 81
DB 516 LKLELV 522

RESULT 20
US-10-873-528-11
Sequence 11, Application US/10873528
Publication No. US2005027681A1
GENERAL INFORMATION:
APPLICANT: Microbial Technics Limited
APPLICANT: Gilbert, Christophe FG
APPLICANT: Hanbro, Philip M
TITLE OF INVENTION: Proteins
FILE REFERENCE: PMC/P211290
CURRENT APPLICATION NUMBER: US/10/873,528
CURRENT FILING DATE: 2004-06-23
PRIOR APPLICATION NUMBER: US/09/769,787
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: GB 9816337.1
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: US 60/125164
PRIOR FILING DATE: 1999-03-19
NUMBER OF SEQ ID NOS: 388
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 11
LENGTH: 1099
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-10-873-528-11

Query Match 3.2%; Score 7; DB 6; Length 1099;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 TDLYTDL 69
DB 402 TDLYTDL 408

```

RESULT 21
; Sequence 2, Application US/11128572
; Publication No. US20050266513A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim International GmbH
; TITLE OF INVENTION: Method for identifying compounds that have the potential to inhibit
; FILE OF INVENTION: activity of Myc
; FILE REFERENCE: Case 12-0231-PCT
; CURRENT APPLICATION NUMBER: US/11/128, 572
; CURRENT FILING DATE: 2005-05-13
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 4374
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-128-572-2

Query Match      3.2%: Score 7; DB 6; Length 4374;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      75 LKLEEV 81
      |||||
Db      1557 LKLEEV 1563

RESULT 22
US-10-995-561-777
; Sequence 777, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995, 561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 777
; LENGTH: 5335
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-777

Query Match      3.2%: Score 7; DB 6; Length 5335;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      78 LLEVLGG 84
      |||||
Db      23 LLEVLGG 29

RESULT 23
US-11-045-024-2635
; Sequence 2635, Application US/11045024
; Publication No. US20050271676A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Livingston, Brian
; APPLICANT: Cheesnut, Robert
; APPLICANT: Baker, Denise Marie
; APPLICANT: Celis, Eserben
; APPLICANT: Kubo, Ralph
; APPLICANT: Grey, Howard M.
; APPLICANT: Eppimmune Inc.

```

```

TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
FILE REFERENCE: 2060.0040007
CURRENT FILING DATE: 2005-01-28
PRIOR APPLICATION NUMBER: US 09/412,863
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 08/027,146
PRIOR FILING DATE: 1993-03-05
PRIOR APPLICATION NUMBER: US 08/073,205
PRIOR FILING DATE: 1993-06-04
PRIOR APPLICATION NUMBER: US 08/159,184
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/159,339
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/205,713
PRIOR FILING DATE: 1994-03-04
NUMBER OF SEQ ID NOS: 14528
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2635
LENGTH: 8
TYPE: FRT
ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-2635

Query Match      2.7%  Score 6;  DB 7;  Length 8;
Best Local Similarity 100.0%;  Pred. No.1e+05;
Matches      6;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;

QY      134  TLELLE 139
      |||||
Db      1  TLELLE 6

RESULT 24
US-11-045-024-5011
Sequence 5011, Application US/11045024
Publication No. US20050271676A1
GENERAL INFORMATION:
APPLICANT: Sette, Alessandro
APPLICANT: Sidney, John
APPLICANT: Southwood, Scott
APPLICANT: Livingston, Brian
APPLICANT: Chesnut, Robert
APPLICANT: Baker, Denise Marie
APPLICANT: Celis, Bastejan
APPLICANT: Kubo, Ralph
APPLICANT: Grey, Howard M.
APPLICANT: Epimmune Inc.
TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
FILE REFERENCE: 2060.0040007
CURRENT FILING DATE: 2005-01-28
PRIOR APPLICATION NUMBER: US/11/045,024
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 09/412,863
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 08/027,146
PRIOR FILING DATE: 1993-03-05
PRIOR APPLICATION NUMBER: US 08/073,205
PRIOR FILING DATE: 1993-06-04
PRIOR APPLICATION NUMBER: US 08/103,396
PRIOR FILING DATE: 1993-08-06
PRIOR APPLICATION NUMBER: US 08/159,184
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/159,339
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/205,713
PRIOR FILING DATE: 1994-03-04
PRIOR APPLICATION NUMBER: US 08/347,610

```

;; PRIOR FILING DATE: 1994-12-01
;; NUMBER OF SEQ ID NOS: 14528
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 5011
;; LENGTH: 8
;; TYPE: PRT
;; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-5011

Query Match 2.7%; Score 6; DB 7; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLEBLE 139
|||
Db 1 TLEBLE 6

RESULT 25
US-11-045-024-2649
; Sequence 2649, Application US/11045024
; Publication No. US20050271676A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert
; APPLICANT: Baker, Denise Marie
; APPLICANT: Celis, Estebean
; APPLICANT: Kubo, Ralph
; APPLICANT: Grey, Howard M.
; APPLICANT: Eptimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
; FILE REFERENCE: 2060.0040007
; CURRENT APPLICATION NUMBER: US/11/045,024
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 09/412,863
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: US 08/027,146
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: US 08/073,205
; PRIOR FILING DATE: 1993-06-04
; PRIOR APPLICATION NUMBER: US 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: US 08/159,184
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; NUMBER OF SEQ ID NOS: 14528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2649
; LENGTH: 9
; TYPE: PRT
; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-2649

Query Match 2.7%; Score 6; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLEBLE 139
|||
Db 2 TLEBLE 7

RESULT 26
US-11-045-024-3576

; Sequence 3576, Application US/11045024
; Publication No. US20050271676A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert
; APPLICANT: Baker, Denise Marie
; APPLICANT: Celis, Estebean
; APPLICANT: Kubo, Ralph
; APPLICANT: Grey, Howard M.
; APPLICANT: Eptimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
; FILE REFERENCE: 2060.0040007
; CURRENT APPLICATION NUMBER: US/11/045,024
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 09/412,863
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: US 08/027,146
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: US 08/073,205
; PRIOR FILING DATE: 1993-06-04
; PRIOR APPLICATION NUMBER: US 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: US 08/159,184
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; NUMBER OF SEQ ID NOS: 14528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3576
; LENGTH: 9
; TYPE: PRT
; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-3576

Query Match 2.7%; Score 6; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLEBLE 139
|||
Db 1 TLEBLE 6

RESULT 27
US-11-045-024-5023
; Sequence 5023, Application US/11045024
; Publication No. US20050271676A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert
; APPLICANT: Baker, Denise Marie
; APPLICANT: Celis, Estebean
; APPLICANT: Kubo, Ralph
; APPLICANT: Grey, Howard M.
; APPLICANT: Eptimmune Inc.
; TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
; FILE REFERENCE: 2060.0040007
; CURRENT APPLICATION NUMBER: US/11/045,024
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 09/412,863
; PRIOR FILING DATE: 1999-10-05


```

; PRIOR APPLICATION NUMBER: US 08/027,146
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: US 08/073,205
; PRIOR FILING DATE: 1993-06-04
; PRIOR APPLICATION NUMBER: US 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: US 08/159,184
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; NUMBER OF SEQ ID NOS: 14528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5023
; LENGTH: 9
; TYPE: PRT
; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-5023

Query Match
Best Local Similarity 100.0%; Score 6; DB 7; Length 9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLELLE 139
DB 2 TLELLE 7

RESULT 28
US-11-045-024-6897
; Sequence 6897, Application US/11045024
; Publication No. US20050271676A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert
; APPLICANT: Baker, Denise Marie
; APPLICANT: Celis, Esceban
; APPLICANT: Kubo, Ralph
; APPLICANT: Grey, Howard M.
; APPLICANT: EpiImmune Inc.
; TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
; FILE REFERENCE: 2060.0040007
; CURRENT APPLICATION NUMBER: US/11/045,024
; PRIOR FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 09/412,863
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: US 08/027,146
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: US 08/073,205
; PRIOR FILING DATE: 1993-06-04
; PRIOR APPLICATION NUMBER: US 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: US 08/159,184
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; NUMBER OF SEQ ID NOS: 14528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6897
; LENGTH: 9
; TYPE: PRT
; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
```

```

US-11-045-024-6897

Query Match
Best Local Similarity 100.0%; Score 6; DB 7; Length 9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLELLE 139
DB 2 TLELLE 7

RESULT 29
US-11-045-024-10720
; Sequence 10720, Application US/11045024
; Publication No. US20050271676A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert
; APPLICANT: Baker, Denise Marie
; APPLICANT: Celis, Esceban
; APPLICANT: Kubo, Ralph
; APPLICANT: Grey, Howard M.
; APPLICANT: EpiImmune Inc.
; TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
; FILE REFERENCE: 2060.0040007
; CURRENT APPLICATION NUMBER: US/11/045,024
; PRIOR FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 09/412,863
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: US 08/027,146
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: US 08/073,205
; PRIOR FILING DATE: 1993-06-04
; PRIOR APPLICATION NUMBER: US 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: US 08/159,184
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; NUMBER OF SEQ ID NOS: 14528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10720
; LENGTH: 9
; TYPE: PRT
; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-10720

Query Match
Best Local Similarity 100.0%; Score 6; DB 7; Length 9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLELLE 139
DB 1 TLELLE 6

RESULT 30
US-11-045-024-12496
; Sequence 12496, Application US/11045024
; Publication No. US20050271676A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Livingston, Brian
```

APPLICANT: Chesnut, Robert
APPLICANT: Baker, Denise Marie
APPLICANT: Celis, Betteban
APPLICANT: Kubo, Ralph
APPLICANT: Grey, Howard M.
APPLICANT: EpiImmune Inc.
TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
FILE REFERENCE: 2060.0040007
CURRENT APPLICATION NUMBER: US/11/045,024
CURRENT FILING DATE: 2005-01-28
PRIOR APPLICATION NUMBER: US 09/412,863
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 08/027,146
PRIOR FILING DATE: 1993-03-05
PRIOR APPLICATION NUMBER: US 08/073,205
PRIOR FILING DATE: 1993-06-04
PRIOR APPLICATION NUMBER: US 08/103,396
PRIOR FILING DATE: 1993-08-06
PRIOR APPLICATION NUMBER: US 08/159,184
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/159,339
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/205,713
PRIOR FILING DATE: 1994-03-04
PRIOR APPLICATION NUMBER: US 08/347,610
PRIOR FILING DATE: 1994-12-01
NUMBER OF SEQ ID NOS: 14528
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12496
LENGTH: 9
TYPE: PRT
ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-12496

Query Match 2.7%; Score 6; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLELLE 139
Db 1 TLELLE 6

RESULT 31
US-11-045-024-14001
Sequence 14001, Application US/11045024
Publication No. US20050271676A1
GENERAL INFORMATION:
APPLICANT: Sette, Alessandro
APPLICANT: Sidney, John
APPLICANT: Southwood, Scott
APPLICANT: Livingston, Brian
APPLICANT: Chesnut, Robert
APPLICANT: Baker, Denise Marie
APPLICANT: Celis, Betteban
APPLICANT: Kubo, Ralph
APPLICANT: Grey, Howard M.
APPLICANT: EpiImmune Inc.
TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
FILE REFERENCE: 2060.0040007
CURRENT APPLICATION NUMBER: US/11/045,024
CURRENT FILING DATE: 2005-01-28
PRIOR APPLICATION NUMBER: US 09/412,863
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 08/027,146
PRIOR FILING DATE: 1993-03-05
PRIOR APPLICATION NUMBER: US 08/073,205
PRIOR FILING DATE: 1993-06-04
PRIOR APPLICATION NUMBER: US 08/103,396
PRIOR FILING DATE: 1993-08-06
PRIOR APPLICATION NUMBER: US 08/159,184

PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/159,339
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/205,713
PRIOR FILING DATE: 1994-03-04
PRIOR APPLICATION NUMBER: US 08/347,610
PRIOR FILING DATE: 1994-12-01
NUMBER OF SEQ ID NOS: 14528
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14001
LENGTH: 9
TYPE: PRT
ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-14001

Query Match 2.7%; Score 6; DB 7; Length 9;
Best Local Similarity 100.0%; Pred. No. 1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLELLE 139
Db 2 TLELLE 7

RESULT 32
US-11-045-024-3582
Sequence 3582, Application US/11045024
Publication No. US20050271676A1
GENERAL INFORMATION:
APPLICANT: Sette, Alessandro
APPLICANT: Sidney, John
APPLICANT: Southwood, Scott
APPLICANT: Livingston, Brian
APPLICANT: Chesnut, Robert
APPLICANT: Baker, Denise Marie
APPLICANT: Celis, Betteban
APPLICANT: Kubo, Ralph
APPLICANT: Grey, Howard M.
APPLICANT: EpiImmune Inc.
TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
FILE REFERENCE: 2060.0040007
CURRENT APPLICATION NUMBER: US/11/045,024
CURRENT FILING DATE: 2005-01-28
PRIOR APPLICATION NUMBER: US 09/412,863
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 08/027,146
PRIOR FILING DATE: 1993-03-05
PRIOR APPLICATION NUMBER: US 08/073,205
PRIOR FILING DATE: 1993-06-04
PRIOR APPLICATION NUMBER: US 08/103,396
PRIOR FILING DATE: 1993-08-06
PRIOR APPLICATION NUMBER: US 08/159,184
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/205,713
PRIOR FILING DATE: 1994-03-04
PRIOR APPLICATION NUMBER: US 08/347,610
PRIOR FILING DATE: 1994-12-01
NUMBER OF SEQ ID NOS: 14528
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3582
LENGTH: 10
TYPE: PRT
ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-3582

Query Match 2.7%; Score 6; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLELLE 139

Db 2 TLELLE 7

```
RESULT 33
US-11-045-024-5043
; Sequence 5043, Application US/11045024
; Publication No. US20050271676A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert
; APPLICANT: Baker, Denise Marie
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph
; APPLICANT: Grey, Howard M.
; APPLICANT: EpiImmune Inc.
; TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
; FILE REFERENCE: 2060.0040007
; CURRENT APPLICATION NUMBER: US/11/045,024
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 09/412,863
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: US 08/027,146
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: US 08/073,205
; PRIOR FILING DATE: 1993-06-04
; PRIOR APPLICATION NUMBER: US 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: US 08/159,184
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
; NUMBER OF SEQ ID NOS: 14528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5043
; LENGTH: 10
; TYPE: PRT
; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-5043

Query Match 2.7% Score 6; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLELLE 139
Db 3 TLELLE 8

RESULT 34
US-11-045-024-10716
; Sequence 10716, Application US/11045024
; Publication No. US20050271676A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert
; APPLICANT: Baker, Denise Marie
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph
; APPLICANT: Grey, Howard M.
; APPLICANT: EpiImmune Inc.
; TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
```

```
; TITLE OF INVENTION: Virus-1 Using Peptide and Nucleic Acid Compositions
; FILE REFERENCE: 2060.0040007
; CURRENT APPLICATION NUMBER: US/11/045,024
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 09/412,863
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: US 08/027,146
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: US 08/073,205
; PRIOR FILING DATE: 1993-06-04
; PRIOR APPLICATION NUMBER: US 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: US 08/159,184
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/347,610
; NUMBER OF SEQ ID NOS: 14528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10716
; LENGTH: 10
; TYPE: PRT
; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-10716

Query Match 2.7% Score 6; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 TLELLE 139
Db 2 TLELLE 7

RESULT 35
US-11-045-024-12492
; Sequence 12492, Application US/11045024
; Publication No. US20050271676A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert
; APPLICANT: Baker, Denise Marie
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph
; APPLICANT: Grey, Howard M.
; APPLICANT: EpiImmune Inc.
; TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
; FILE REFERENCE: 2060.0040007
; CURRENT APPLICATION NUMBER: US/11/045,024
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 09/412,863
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: US 08/027,146
; PRIOR FILING DATE: 1993-03-05
; PRIOR APPLICATION NUMBER: US 08/073,205
; PRIOR FILING DATE: 1993-06-04
; PRIOR APPLICATION NUMBER: US 08/103,396
; PRIOR FILING DATE: 1993-08-06
; PRIOR APPLICATION NUMBER: US 08/159,184
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/159,339
; PRIOR FILING DATE: 1993-11-29
; PRIOR APPLICATION NUMBER: US 08/205,713
; PRIOR FILING DATE: 1994-03-04
; PRIOR APPLICATION NUMBER: US 08/347,610
; PRIOR FILING DATE: 1994-12-01
```

NUMBER OF SEQ ID NOS: 14528
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12492
LENGTH: 10
TYPE: PRT
ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-12492

Query Match 2.7%; Score 6; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 134 TLEFILE 139
Db 2 TLEFILE 7

RESULT 36
US-11-045-024-12853
Sequence 12853, Application US/11045024
Publication No. US20050271676A1
GENERAL INFORMATION:

APPLICANT: Sette, Alessandro
APPLICANT: Sidney, John
APPLICANT: Southwood, Scott
APPLICANT: Livingston, Brian
APPLICANT: Chesnut, Robert
APPLICANT: Baker, Denise Marie
APPLICANT: Cells, Estebean
APPLICANT: Kubo, Ralph
APPLICANT: Grey, Howard M.
APPLICANT: EpiImmune Inc.
TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
Virus-1 Using Peptide and Nucleic Acid Compositions
FILE REFERENCE: 2060.0040007
CURRENT FILING DATE: 2005-01-28
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 09/412,863
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 08/027,146
PRIOR FILING DATE: 1993-03-05
PRIOR APPLICATION NUMBER: US 08/073,205
PRIOR FILING DATE: 1993-06-04
PRIOR APPLICATION NUMBER: US 08/103,396
PRIOR FILING DATE: 1993-08-06
PRIOR APPLICATION NUMBER: US 08/159,184
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/159,339
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/205,713
PRIOR FILING DATE: 1994-03-04
PRIOR APPLICATION NUMBER: US 08/347,610
PRIOR FILING DATE: 1994-12-01
NUMBER OF SEQ ID NOS: 14528
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12853
LENGTH: 10
TYPE: PRT
ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-12853

Query Match 2.7%; Score 6; DB 7; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 134 TLEFILE 139
Db 3 TLEFILE 8

RESULT 37
US-11-045-024-13345
Sequence 13345, Application US/11045024

Publication No. US20050271676A1
GENERAL INFORMATION:

APPLICANT: Sette, Alessandro
APPLICANT: Sidney, John
APPLICANT: Southwood, Scott
APPLICANT: Livingston, Brian
APPLICANT: Chesnut, Robert
APPLICANT: Baker, Denise Marie
APPLICANT: Cells, Estebean
APPLICANT: Kubo, Ralph
APPLICANT: Grey, Howard M.
APPLICANT: EpiImmune Inc.

TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
Virus-1 Using Peptide and Nucleic Acid Compositions
FILE REFERENCE: 2060.0040007
CURRENT FILING DATE: 2005-01-28
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 09/412,863
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 08/027,146
PRIOR FILING DATE: 1993-03-05
PRIOR APPLICATION NUMBER: US 08/073,205
PRIOR FILING DATE: 1993-06-04
PRIOR APPLICATION NUMBER: US 08/103,396
PRIOR FILING DATE: 1993-08-06
PRIOR APPLICATION NUMBER: US 08/159,184
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/159,339
PRIOR FILING DATE: 1993-11-29
PRIOR APPLICATION NUMBER: US 08/205,713
PRIOR FILING DATE: 1994-03-04
PRIOR APPLICATION NUMBER: US 08/347,610
PRIOR FILING DATE: 1994-12-01
NUMBER OF SEQ ID NOS: 14528
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13345
LENGTH: 15
TYPE: PRT
ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-13345

Query Match 2.7%; Score 6; DB 7; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 134 TLEFILE 139
Db 5 TLEFILE 10

RESULT 38
US-11-045-024-13348
Sequence 13348, Application US/11045024
Publication No. US20050271676A1
GENERAL INFORMATION:

APPLICANT: Sette, Alessandro
APPLICANT: Sidney, John
APPLICANT: Southwood, Scott
APPLICANT: Livingston, Brian
APPLICANT: Chesnut, Robert
APPLICANT: Baker, Denise Marie
APPLICANT: Cells, Estebean
APPLICANT: Kubo, Ralph
APPLICANT: Grey, Howard M.
APPLICANT: EpiImmune Inc.

TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
Virus-1 Using Peptide and Nucleic Acid Compositions
FILE REFERENCE: 2060.0040007
CURRENT FILING DATE: 2005-01-28
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: US 08/027,146

```
;; PRIOR FILING DATE: 1993-03-05
;; PRIOR APPLICATION NUMBER: US 08/073,205
;; PRIOR FILING DATE: 1993-06-04
;; PRIOR APPLICATION NUMBER: US 08/103,396
;; PRIOR FILING DATE: 1993-08-06
;; PRIOR APPLICATION NUMBER: US 08/159,184
;; PRIOR FILING DATE: 1993-11-29
;; PRIOR APPLICATION NUMBER: US 08/159,339
;; PRIOR FILING DATE: 1993-11-29
;; PRIOR APPLICATION NUMBER: US 08/205,713
;; PRIOR FILING DATE: 1994-03-04
;; PRIOR APPLICATION NUMBER: US 08/347,610
;; PRIOR FILING DATE: 1994-12-01
;; NUMBER OF SEQ ID NOS: 14528
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 13348
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-13348

Query Match          2.7%; Score 6; DB 7; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      134 TLELLE 139
DB      8 TLELLE 13

RESULT 39
US-11-045-024-13515
;; Sequence 13515, Application US/11045024
;; Publication No. US20050271676A1
;; GENERAL INFORMATION:
;; APPLICANT: Sette, Alessandro
;; APPLICANT: Sidney, John
;; APPLICANT: Southwood, Scott
;; APPLICANT: Livingston, Brian
;; APPLICANT: Chesnut, Robert
;; APPLICANT: Baker, Denise Marie
;; APPLICANT: Celis, Steban
;; APPLICANT: Kubo, Ralph
;; APPLICANT: Grey, Howard M.
;; APPLICANT: EpiImmune Inc.
;; TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
;; FILE REFERENCE: 2060.0040007
;; CURRENT APPLICATION NUMBER: US/11/045,024
;; CURRENT FILING DATE: 2005-01-28
;; PRIOR APPLICATION NUMBER: US 09/412,863
;; PRIOR FILING DATE: 1999-10-05
;; PRIOR APPLICATION NUMBER: US 08/027,146
;; PRIOR FILING DATE: 1993-03-05
;; PRIOR APPLICATION NUMBER: US 08/073,205
;; PRIOR FILING DATE: 1993-06-04
;; PRIOR APPLICATION NUMBER: US 08/103,396
;; PRIOR FILING DATE: 1993-08-06
;; PRIOR APPLICATION NUMBER: US 08/159,184
;; PRIOR FILING DATE: 1993-11-29
;; PRIOR APPLICATION NUMBER: US 08/159,339
;; PRIOR FILING DATE: 1993-11-29
;; PRIOR APPLICATION NUMBER: US 08/205,713
;; PRIOR FILING DATE: 1994-03-04
;; PRIOR APPLICATION NUMBER: US 08/347,610
;; PRIOR FILING DATE: 1994-12-01
;; NUMBER OF SEQ ID NOS: 14528
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 13515
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-13515
```

```
Query Match          2.7%; Score 6; DB 7; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      134 TLELLE 139
DB      1 TLELLE 6

RESULT 40
US-11-045-024-13516
;; Sequence 13516, Application US/11045024
;; Publication No. US20050271676A1
;; GENERAL INFORMATION:
;; APPLICANT: Sette, Alessandro
;; APPLICANT: Sidney, John
;; APPLICANT: Southwood, Scott
;; APPLICANT: Livingston, Brian
;; APPLICANT: Chesnut, Robert
;; APPLICANT: Baker, Denise Marie
;; APPLICANT: Celis, Steban
;; APPLICANT: Kubo, Ralph
;; APPLICANT: Grey, Howard M.
;; APPLICANT: EpiImmune Inc.
;; TITLE OF INVENTION: Inducing Cellular Responses to Human Immunodeficiency
;; FILE REFERENCE: 2060.0040007
;; CURRENT APPLICATION NUMBER: US/11/045,024
;; CURRENT FILING DATE: 2005-01-28
;; PRIOR APPLICATION NUMBER: US 09/412,863
;; PRIOR FILING DATE: 1999-10-05
;; PRIOR APPLICATION NUMBER: US 08/027,146
;; PRIOR FILING DATE: 1993-03-05
;; PRIOR APPLICATION NUMBER: US 08/073,205
;; PRIOR FILING DATE: 1993-06-04
;; PRIOR APPLICATION NUMBER: US 08/103,396
;; PRIOR FILING DATE: 1993-08-06
;; PRIOR APPLICATION NUMBER: US 08/159,184
;; PRIOR FILING DATE: 1993-11-29
;; PRIOR APPLICATION NUMBER: US 08/159,339
;; PRIOR FILING DATE: 1993-11-29
;; PRIOR APPLICATION NUMBER: US 08/205,713
;; PRIOR FILING DATE: 1994-03-04
;; PRIOR APPLICATION NUMBER: US 08/347,610
;; PRIOR FILING DATE: 1994-12-01
;; NUMBER OF SEQ ID NOS: 14528
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO: 13516
;; LENGTH: 15
;; TYPE: PRT
;; ORGANISM: HUMAN IMMUNODEFICIENCY VIRUS
US-11-045-024-13516

Query Match          2.7%; Score 6; DB 7; Length 15;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      134 TLELLE 139
DB      1 TLELLE 6

RESULT 41
US-11-096-725-28
;; Sequence 28, Application US/11096725
;; Publication No. US20060013820A1
;; GENERAL INFORMATION:
;; APPLICANT: BONNET, Dominique
;; APPLICANT: BROWN, Carlton B
;; APPLICANT: GEORGES, Bertrand
;; APPLICANT: SIZER, Philip J
;; TITLE OF INVENTION: Antigen Delivery Vectors and Constructs
US-11-045-024-13515
```

```
/ FILE REFERENCE: KLB-001
/ CURRENT APPLICATION NUMBER: US/11/096,725
/ CURRENT FILING DATE: 2005-04-01
/ PRIOR APPLICATION NUMBER: GBO408164.2
/ PRIOR FILING DATE: 2004-04-13
/ NUMBER OF SEQ ID NOS: 40
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 28
/ LENGTH: 35
/ TYPE: PRT
/ ORGANISM: Human immunodeficiency virus type 1
US-11-096-725-28

Query Match      2.7%; Score 6; DB 7; Length 35;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      134 TLELLE 139
      |||||
      4 TLELLE 9

RESULT 42
US-11-107-096-78
/ Sequence 78, Application US/11/107096
/ Publication No. US20060003348A1
/ GENERAL INFORMATION:
/ APPLICANT: SIDHU, SACHDEV S.
/ APPLICANT: ZHANG, YINGNAN
/ TITLE OF INVENTION: OMT PDZ MODULATORS
/ FILE REFERENCE: P2100R1
/ CURRENT APPLICATION NUMBER: US/11/107,096
/ CURRENT FILING DATE: 2005-04-15
/ PRIOR APPLICATION NUMBER: US 60/563,157
/ PRIOR FILING DATE: 2004-04-16
/ NUMBER OF SEQ ID NOS: 79
/ SEQ ID NO 78
/ LENGTH: 100
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-107-096-78

Query Match      2.7%; Score 6; DB 7; Length 100;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      137 LLEVR 142
      |||||
      80 LLEVR 85

RESULT 43
US-11-053-076-59
/ Sequence 59, Application US/11/053076
/ Publication No. US20050255460A1
/ GENERAL INFORMATION:
/ APPLICANT: Lu, Peter S.
/ APPLICANT: Schweizer, Johannes
/ APPLICANT: Somora Diaz-Sarmiento, Chamorro
/ APPLICANT: Belmares, Michael P.
/ TITLE OF INVENTION: METHODS OF DIAGNOSING CERVICAL CANCER
/ FILE REFERENCE: VITA-008CIP
/ CURRENT APPLICATION NUMBER: US/11/053,076
/ CURRENT FILING DATE: 2005-02-07
/ PRIOR APPLICATION NUMBER: PCT/US03/28508
/ PRIOR FILING DATE: 2003-09-09
/ PRIOR APPLICATION NUMBER: 10/630,590
/ PRIOR FILING DATE: 2003-07-29
/ PRIOR APPLICATION NUMBER: 60/490,094
/ PRIOR FILING DATE: 2003-07-25
/ PRIOR APPLICATION NUMBER: 60/450,464
/ PRIOR FILING DATE: 2003-02-27
/ PRIOR APPLICATION NUMBER: 60/409,298
```

```
/ PRIOR FILING DATE: 2002-09-09
/ PRIOR APPLICATION NUMBER: 10/630,590
/ PRIOR FILING DATE: 2003-07-29
/ PRIOR APPLICATION NUMBER: PCT/US02/24655
/ PRIOR FILING DATE: 2002-08-02
/ PRIOR APPLICATION NUMBER: 60/309,841
/ PRIOR FILING DATE: 2001-08-03
/ PRIOR APPLICATION NUMBER: 60/360,061
/ PRIOR FILING DATE: 2002-02-25
/ PRIOR APPLICATION NUMBER: 10/080,273
/ PRIOR FILING DATE: 2002-02-19
/ Removing Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 330
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 59
/ LENGTH: 109
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-11-053-076-59

Query Match      2.7%; Score 6; DB 7; Length 109;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      137 LLEVR 142
      |||||
      93 LLEVR 98

RESULT 44
US-10-485-788A-724
/ Sequence 724, Application US/10/485788A
/ Publication No. US20050282743A1
/ GENERAL INFORMATION:
/ APPLICANT: Rabinowitz, Joshua D.
/ APPLICANT: Schweizer, Johannes
/ APPLICANT: Carrick, Deanna Marie
/ APPLICANT: Arbor Vita Corporation
/ TITLE OF INVENTION: Molecular Interactions in Cells
/ FILE REFERENCE: 20054-00320US
/ CURRENT APPLICATION NUMBER: US/10/485,788A
/ CURRENT FILING DATE: 2004-02-03
/ PRIOR APPLICATION NUMBER: US 60/309,841
/ PRIOR FILING DATE: 2001-08-03
/ PRIOR APPLICATION NUMBER: US 60/360,061
/ PRIOR FILING DATE: 2002-02-25
/ PRIOR APPLICATION NUMBER: WO PCT/US02/24655
/ PRIOR FILING DATE: 2002-08-02
/ NUMBER OF SEQ ID NOS: 841
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 724
/ LENGTH: 113
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-485-788A-724

Query Match      2.7%; Score 6; DB 6; Length 113;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      77 KLEVL 82
      |||||
      22 KLEVL 27

RESULT 45
US-11-053-076-94
/ Sequence 94, Application US/11/053076
/ Publication No. US20050255460A1
/ GENERAL INFORMATION:
/ APPLICANT: Lu, Peter S.
/ APPLICANT: Schweizer, Johannes
```

APPLICANT: Somoza Diaz-Sarmiento, Chamorro
APPLICANT: Belmares, Michael P.
TITLE OF INVENTION: METHODS OF DIAGNOSING CERVICAL CANCER
FILE REFERENCE: VITA-008CIP
CURRENT FILING DATE: 2005-02-07
PRIOR APPLICATION NUMBER: PCT/US03/28508
PRIOR FILING DATE: 2003-09-09
PRIOR APPLICATION NUMBER: 10/630,590
PRIOR FILING DATE: 2003-07-29
PRIOR APPLICATION NUMBER: 60/490,094
PRIOR FILING DATE: 2003-07-25
PRIOR APPLICATION NUMBER: 60/450,464
PRIOR FILING DATE: 2003-02-27
PRIOR APPLICATION NUMBER: 60/409,298
PRIOR FILING DATE: 2002-09-09
PRIOR APPLICATION NUMBER: 10/630,590
PRIOR FILING DATE: 2003-07-29
PRIOR APPLICATION NUMBER: 60/360,061
PRIOR FILING DATE: 2002-02-25
PRIOR APPLICATION NUMBER: 10/080,273
PRIOR FILING DATE: 2002-02-19
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 330
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 94
LENGTH: 113
TYPE: PRT
ORGANISM: Homo sapiens
US-11-053-076-94

Query Match
Best Local Similarity 100.0%; Score 6; DB 7; Length 113;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 KLEVL 82
Db 22 KLEVL 27

RESULT 46
US-11-055-822-146
Sequence 146, Application US/11055822
Publication No. US20050260707A1
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Krogger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
FILE REFERENCE: BGI-121CPCN
CURRENT APPLICATION NUMBER: US/11/055,822
CURRENT FILING DATE: 2005-02-11
PRIOR APPLICATION NUMBER: 09/606,740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/142,101
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 60/148,613
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/187,970
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: DE 19930476.9
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: DE 19931415.2
PRIOR FILING DATE: 1999-07-06

PRIOR APPLICATION NUMBER: DE 19931418.7
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1158
SEQ ID NO 146
LENGTH: 126
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-11-055-822-146

Query Match
Best Local Similarity 100.0%; Score 6; DB 7; Length 126;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 LLEVR 142
Db 58 LLEVR 63

RESULT 47
US-11-052-554A-88
Sequence 88, Application US/11052554A
Publication No. US2005028866A1
GENERAL INFORMATION:
APPLICANT: Sachdeva, et al.
TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
FILE REFERENCE: 30853/40359A
CURRENT APPLICATION NUMBER: US/11/052,554A
CURRENT FILING DATE: 2005-02-07
PRIOR APPLICATION NUMBER: US 60/589,227
PRIOR FILING DATE: 2004-07-20
PRIOR APPLICATION NUMBER: IN 173/DEL/2004
PRIOR FILING DATE: 2004-02-06
NUMBER OF SEQ ID NOS: 763
SOFTWARE: Patent version 3.3
SEQ ID NO 88
LENGTH: 134
TYPE: PRT
ORGANISM: Haemophilus influenzae Rd
US-11-052-554A-88

Query Match
Best Local Similarity 100.0%; Score 6; DB 7; Length 134;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 VDKALQ 108
Db 36 VDKALQ 41

RESULT 48
US-11-087-099-7788
Sequence 7788, Application US/11087099
Publication No. US20060041961A1
GENERAL INFORMATION:
APPLICANT: Abad, Mark S. et al.
TITLE OF INVENTION: Genes and Uses for Plant Improvement
FILE REFERENCE: 38-21(53450)B EP
CURRENT APPLICATION NUMBER: US/11/087,099
CURRENT FILING DATE: 2005-03-22
NUMBER OF SEQ ID NOS: 12464
SEQ ID NO 7788
LENGTH: 145
TYPE: PRT
ORGANISM: Zea mays
US-11-087-099-7788

Query Match
Best Local Similarity 100.0%; Score 6; DB 7; Length 145;

Best Local Similarity 100.0%; Pred. No. 84;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 143 QOEER 148
|||||

Db 130 QOEER 135

RESULT 49

US-10-510-386-194
; Sequence 194, Application US/10510386
; Publication No. US20050244922A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Clausen, Ib Groth
; APPLICANT: Jorgensen, Steen Troels
; APPLICANT: Olsen, Peter Bjarke
; APPLICANT: Rasmussen, Michael Dolberg
; TITLE OF INVENTION: Improved Bacillus Host Cell
; FILE REFERENCE: 10294.204-US
; CURRENT APPLICATION NUMBER: US/10/510,386
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 248
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 194
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Bacillus licheniformis
US-10-510-386-194

Query Match

Best Local Similarity 100.0%; Score 6; DB 6; Length 163;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 TKVNS 53
|||||

Db 27 TKVNS 32

RESULT 50

US-11-087-099-5484
; Sequence 5484, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 5484
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Trilicium aestivum
US-11-087-099-5484

Query Match

Best Local Similarity 100.0%; Score 6; DB 7; Length 165;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 ISGPLS 14
|||||

Db 104 ISGPLS 109

Search completed: March 4, 2006, 07:14:41
Job time : 22 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OW protein - protein search, using sw model

Run on: March 4, 2006, 07:10:46 ; Search time 165 Seconds
(without alignments)
557.106 Million cell updates/sec

Title: US-10-695-994a-7

Perfect score: 220

Sequence: 1 MELORTSISGPISPAYNGQ.....QNYKNFNRRTADSHSMSCM 220

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1867569 seqs, 417829326 residues

Word size: 6

Total number of hits satisfying chosen parameters: 4362

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Published Applications AA Main:

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	220	100.0	220	US-10-695-994-7	Sequence 7, Appli
2	124	56.4	2154	US-10-695-994-4	Sequence 4, Appli
3	124	56.4	2154	US-10-805-684-150	Sequence 150, App
4	97	44.1	1630	US-10-408-765A-2431	Sequence 2431, Ap
5	97	44.1	1630	US-10-170-385-15	Sequence 15, Appl
6	38	17.3	57	US-09-945-917-43	Sequence 43, Appl
7	38	17.3	2106	US-10-408-765A-2093	Sequence 2093, Ap
8	38	17.3	2137	US-10-756-149-4806	Sequence 4806, Ap
9	27	12.3	2388	US-10-972-637-2	Sequence 2, Appli
10	27	12.3	2390	US-10-756-149-5749	Sequence 5749, Ap
11	25	11.4	51	US-09-945-917-40	Sequence 40, Appl
12	23	10.5	806	US-10-264-049-2841	Sequence 2841, Ap
13	22	10.0	2559	US-10-756-149-4730	Sequence 4730, Ap
14	19	8.6	2291	US-11-097-143-12420	Sequence 12420, A
15	16	7.3	48	US-09-945-917-46	Sequence 46, Appl
16	15	6.8	17	US-10-862-195-1550	Sequence 1550, Ap
17	15	6.8	17	US-10-862-195-1550	Sequence 1551, Ap
18	14	6.4	14	US-10-695-994-20	Sequence 20, Appl
19	14	6.4	1355	US-10-369-493-6374	Sequence 6374, Ap
20	11	5.0	1355	US-10-369-493-5251	Sequence 5251, Ap
21	11	5.0	1355	US-10-369-493-5252	Sequence 5252, Ap
22	11	5.0	5373	US-10-408-765A-741	Sequence 741, App
23	11	5.0	5430	US-10-805-684-151	Sequence 151, App
24	11	5.0	7201	US-11-097-143-40200	Sequence 40200, A
25	11	4.5	5385	US-11-097-143-26253	Sequence 26253, A
26	10	4.5	5496	US-11-097-143-28275	Sequence 28275, A
27	9	4.1	67	US-10-363-616-413	Sequence 413, App

28	9	4.1	117	5	US-10-450-763-42286	Sequence 42286, A
29	9	4.1	123	4	US-10-425-115-352183	Sequence 352183, A
30	9	4.1	102	4	US-10-424-599-237993	Sequence 237993, A
31	9	4.1	202	4	US-10-425-114-60052	Sequence 60052, A
32	9	4.1	202	4	US-10-425-114-65298	Sequence 65298, A
33	9	4.1	240	4	US-10-425-114-56959	Sequence 56959, A
34	9	4.1	3433	4	US-10-408-765A-731	Sequence 731, App
35	9	4.1	3433	4	US-10-408-765A-732	Sequence 732, App
36	9	4.1	3433	5	US-10-756-149-5681	Sequence 5681, App
37	9	4.1	65	4	US-10-767-701-80379	Sequence 80379, A
38	8	3.6	177	4	US-10-425-115-368921	Sequence 368921, A
39	8	3.6	209	4	US-10-425-115-196724	Sequence 196724, A
40	8	3.6	213	4	US-10-437-963-196932	Sequence 196932, A
41	8	3.6	234	5	US-10-732-923-18430	Sequence 18430, A
42	8	3.6	246	4	US-10-437-963-18843	Sequence 18843, A
43	8	3.6	358	4	US-10-128-714-3119	Sequence 3119, App
44	8	3.6	369	4	US-10-437-963-140721	Sequence 140721, A
45	8	3.6	382	4	US-10-425-115-242617	Sequence 242617, A
46	8	3.6	517	4	US-10-128-714-8119	Sequence 8119, App
47	8	3.6	547	6	US-11-097-143-16161	Sequence 16161, App
48	8	3.6	690	6	US-11-097-143-16218	Sequence 16218, A
49	8	3.6	911	5	US-10-733-969A-5	Sequence 5, Appli
50	8	3.6	1190	4	US-10-369-493-18546	Sequence 18546, A
51	8	3.6	1519	6	US-11-097-143-26697	Sequence 26697, A
52	8	3.6	1605	5	US-10-732-923-18432	Sequence 18432, A
53	8	3.6	2129	5	US-10-450-763-51108	Sequence 51108, A
54	8	3.6	2129	5	US-10-450-763-51108	Sequence 51108, A
55	8	3.6	2156	5	US-10-732-923-18431	Sequence 18431, A
56	8	3.6	2414	5	US-10-473-127-634	Sequence 634, App
57	8	3.6	2414	5	US-10-473-127-641	Sequence 641, App
58	8	3.6	2414	5	US-10-473-127-642	Sequence 642, App
59	8	3.6	2414	5	US-10-473-127-644	Sequence 644, App
60	8	3.6	2414	5	US-10-473-127-646	Sequence 646, App
61	8	3.6	2414	5	US-10-732-923-18449	Sequence 18449, A
62	8	3.6	2414	5	US-10-756-149-5732	Sequence 5732, App
63	8	3.6	4097	6	US-11-097-143-237	Sequence 237, App
64	8	3.6	4101	4	US-10-369-493-6509	Sequence 6509, App
65	8	3.6	4101	4	US-10-369-493-6510	Sequence 6510, App
66	7	3.2	13	4	US-10-695-994-19	Sequence 19, Appli
67	7	3.2	37	4	US-10-424-599-279435	Sequence 279435, A
68	7	3.2	53	4	US-10-437-963-197612	Sequence 197612, A
69	7	3.2	55	4	US-10-029-386-30066	Sequence 30066, A
70	7	3.2	61	4	US-10-425-115-15271	Sequence 15271, A
71	7	3.2	73	4	US-10-425-115-229526	Sequence 229526, A
72	7	3.2	74	4	US-10-279-964-8	Sequence 8, Appli
73	7	3.2	80	4	US-10-425-115-330257	Sequence 330257, A
74	7	3.2	82	4	US-10-425-115-08513	Sequence 208513, A
75	7	3.2	87	4	US-10-424-599-269764	Sequence 269764, A
76	7	3.2	91	4	US-10-408-765A-2732	Sequence 2732, App
77	7	3.2	99	4	US-10-424-599-250552	Sequence 250552, A
78	7	3.2	99	4	US-10-437-963-169460	Sequence 169460, A
79	7	3.2	105	4	US-10-437-963-147630	Sequence 147630, A
80	7	3.2	105	4	US-10-425-115-263525	Sequence 263525, A
81	7	3.2	105	4	US-10-425-115-448629	Sequence 448629, A
82	7	3.2	109	4	US-10-424-599-232699	Sequence 232699, A
83	7	3.2	109	5	US-10-450-763-56850	Sequence 56850, A
84	7	3.2	115	3	US-09-925-288-645	Sequence 645, App
85	7	3.2	115	4	US-10-102-806-645	Sequence 645, App
86	7	3.2	115	4	US-10-156-761-14021	Sequence 14021, A
87	7	3.2	120	3	US-09-864-761-17707	Sequence 17707, A
88	7	3.2	124	4	US-10-437-963-107253	Sequence 107253, A
89	7	3.2	124	4	US-10-767-701-37917	Sequence 37917, A
90	7	3.2	130	4	US-10-424-599-165499	Sequence 165499, A
91	7	3.2	132	4	US-10-276-774-1967	Sequence 1967, App
92	7	3.2	133	3	US-09-870-844-6	Sequence 6, Appli
93	7	3.2	137	4	US-10-424-599-202407	Sequence 202407, A
94	7	3.2	147	4	US-10-437-963-195145	Sequence 195145, A
95	7	3.2	150	4	US-10-424-599-182090	Sequence 182090, A
96	7	3.2	156	4	US-10-363-829-319	Sequence 319, App
97	7	3.2	158	5	US-10-501-282-4406	Sequence 4406, App
98	7	3.2	163	4	US-10-437-963-188861	Sequence 188861, A
99	7	3.2	166	4	US-10-074-557-7	Sequence 7, Appli
100	7	3.2	172	5	US-10-501-282-4408	Sequence 4408, App

101	7	3.2	177	4	US-10-424-599-237731	Sequence 237731,	174	7	3.2	440	4	US-10-424-599-252189	Sequence 252189,
102	7	3.2	178	4	US-10-424-599-191094	Sequence 191094,	175	7	3.2	441	5	US-10-733-923-7313	Sequence 7313, Ap
103	7	3.2	180	4	US-10-309-175-9	Sequence 9, Appl1	176	7	3.2	445	5	US-10-424-599-235922	Sequence 235922,
104	7	3.2	181	4	US-10-425-115-241302	Sequence 241302,	177	7	3.2	450	5	US-10-510-408-119	Sequence 119, App
105	7	3.2	186	4	US-10-104-047-3726	Sequence 3726, Ap	178	7	3.2	453	4	US-10-282-122A-53370	Sequence 53370, A
106	7	3.2	191	5	US-10-732-923-10045	Sequence 10045, A	179	7	3.2	454	5	US-10-450-763-53225	Sequence 53225, A
107	7	3.2	193	4	US-10-425-115-253377	Sequence 253377,	180	7	3.2	470	4	US-10-389-565-1952	Sequence 1952, Ap
108	7	3.2	198	4	US-10-425-115-207330	Sequence 207330,	181	7	3.2	500	4	US-10-032-585-7588	Sequence 7588, Ap
109	7	3.2	199	4	US-10-108-260A-4110	Sequence 4110, Ap	182	7	3.2	524	3	US-09-761-640-9	Sequence 9, Appl1
110	7	3.2	205	4	US-10-260-708-49	Sequence 49, Appl	183	7	3.2	548	4	US-10-369-493-540	Sequence 540, App
111	7	3.2	205	4	US-10-425-115-215289	Sequence 215289,	184	7	3.2	549	4	US-10-104-047-3479	Sequence 3479, Ap
112	7	3.2	211	4	US-10-437-963-174407	Sequence 174407,	185	7	3.2	553	4	US-10-424-599-165507	Sequence 165507,
113	7	3.2	218	4	US-10-029-386-32067	Sequence 32067, A	186	7	3.2	554	5	US-10-830-828-8	Sequence 8, Appl1
114	7	3.2	219	4	US-10-369-493-16530	Sequence 16530, A	187	7	3.2	594	3	US-09-815-242-12107	Sequence 12107, A
115	7	3.2	218	4	US-10-424-599-274042	Sequence 274042,	188	7	3.2	595	5	US-10-282-122A-66841	Sequence 66841, A
116	7	3.2	227	4	US-10-424-599-152023	Sequence 152023,	189	7	3.2	595	4	US-10-369-493-3265	Sequence 3265, Ap
117	7	3.2	230	4	US-10-425-115-344500	Sequence 344500,	190	7	3.2	597	4	US-10-425-115-341910	Sequence 341910,
118	7	3.2	233	4	US-10-156-761-14807	Sequence 14807, A	191	7	3.2	597	4	US-10-437-963-181830	Sequence 181830,
119	7	3.2	246	4	US-10-425-114-42151	Sequence 42151, A	192	7	3.2	598	4	US-10-425-115-320536	Sequence 320536,
120	7	3.2	247	4	US-10-425-115-233921	Sequence 233921,	193	7	3.2	600	4	US-10-092-900A-124	Sequence 124, App
121	7	3.2	250	4	US-10-424-599-165506	Sequence 165506,	194	7	3.2	609	4	US-10-437-963-119462	Sequence 119462,
122	7	3.2	253	4	US-10-369-493-10635	Sequence 10635, A	195	7	3.2	629	4	US-09-867-550-1540	Sequence 1540, Ap
123	7	3.2	256	4	US-10-767-701-40170	Sequence 40170, A	196	7	3.2	642	3	US-10-369-493-13726	Sequence 13726, A
124	7	3.2	259	3	US-09-925-300-953	Sequence 953, App	197	7	3.2	649	4	US-09-862-027-30	Sequence 30, Appl
125	7	3.2	263	4	US-10-425-115-327227	Sequence 327227,	198	7	3.2	656	3	US-10-437-963-175977	Sequence 175977,
126	7	3.2	271	4	US-10-659-755-14	Sequence 14, Appl	199	7	3.2	656	5	US-10-989-228-30	Sequence 30, Appl
127	7	3.2	272	4	US-10-424-599-235920	Sequence 235920,	200	7	3.2	669	4	US-10-282-122A-66822	Sequence 66822, A
128	7	3.2	272	5	US-10-732-923-9974	Sequence 9974, Ap	201	7	3.2	669	4	US-10-282-122A-67679	Sequence 67679, A
129	7	3.2	277	4	US-10-437-963-162937	Sequence 162937,	202	7	3.2	669	4	US-10-104-047-2425	Sequence 2425, Ap
130	7	3.2	280	4	US-10-425-114-52590	Sequence 52590, A	203	7	3.2	673	4	US-10-380-533-133	Sequence 133, App
131	7	3.2	283	4	US-10-425-115-272724	Sequence 272724,	204	7	3.2	690	4	US-10-732-923-3350	Sequence 3350, Ap
132	7	3.2	290	5	US-10-739-930-8391	Sequence 8391, Ap	205	7	3.2	692	5	US-10-732-923-3350	Sequence 3350, Ap
133	7	3.2	291	4	US-10-767-701-43015	Sequence 43015, A	206	7	3.2	698	4	US-10-351-334-200	Sequence 200, App
134	7	3.2	298	4	US-10-425-114-48246	Sequence 48246, A	207	7	3.2	699	5	US-10-785-823-7	Sequence 7, Appl1
135	7	3.2	302	4	US-10-425-114-68383	Sequence 68383, A	208	7	3.2	701	4	US-10-282-122A-66828	Sequence 66828, A
136	7	3.2	304	4	US-10-425-115-344502	Sequence 344502,	209	7	3.2	701	4	US-10-282-122A-69324	Sequence 69324, A
137	7	3.2	306	4	US-10-724-972A-4481	Sequence 4481, Ap	210	7	3.2	702	4	US-10-369-493-4024	Sequence 4024, Ap
138	7	3.2	314	5	US-10-450-763-33610	Sequence 33610, A	211	7	3.2	702	4	US-10-282-122A-67653	Sequence 67653, A
139	7	3.2	316	4	US-10-351-334-239	Sequence 239, App	212	7	3.2	705	4	US-10-389-566-2111	Sequence 2111, Ap
140	7	3.2	321	4	US-10-351-334-355	Sequence 355, App	213	7	3.2	717	4	US-10-369-493-4154	Sequence 4154, Ap
141	7	3.2	324	4	US-10-489-372-18	Sequence 18, Appl	214	7	3.2	730	4	US-10-131-827-8768	Sequence 8768, Ap
142	7	3.2	326	5	US-10-994-726-672	Sequence 672, App	215	7	3.2	731	4	US-10-437-963-174639	Sequence 174639,
143	7	3.2	329	5	US-10-501-282-5654	Sequence 5654, Ap	216	7	3.2	731	4	US-10-108-260A-7451	Sequence 7451, Ap
144	7	3.2	344	4	US-10-425-115-232166	Sequence 232166,	217	7	3.2	731	5	US-10-450-763-47811	Sequence 47811, Ap
145	7	3.2	344	6	US-10-732-923-10148	Sequence 10148, A	218	7	3.2	737	3	US-09-955-732-13	Sequence 13, Appl
146	7	3.2	344	6	US-11-097-143-3432	Sequence 3432, Ap	219	7	3.2	764	3	US-09-740-043-4	Sequence 4, Appl1
147	7	3.2	347	5	US-10-994-726-671	Sequence 671, App	220	7	3.2	764	4	US-10-072-012-543	Sequence 543, App
148	7	3.2	357	5	US-10-732-923-10043	Sequence 10043, A	221	7	3.2	836	4	US-10-437-963-122882	Sequence 122882,
149	7	3.2	357	5	US-10-732-923-10066	Sequence 10066, A	222	7	3.2	854	5	US-10-756-149-5164	Sequence 5164, Ap
150	7	3.2	357	5	US-10-732-923-10066	Sequence 10066, A	223	7	3.2	857	4	US-10-425-114-45539	Sequence 45539, A
151	7	3.2	358	5	US-10-732-923-10067	Sequence 10067, A	224	7	3.2	857	4	US-10-425-115-205907	Sequence 205907,
152	7	3.2	358	6	US-11-069-633-45	Sequence 45, Appl	225	7	3.2	861	6	US-11-097-143-15876	Sequence 15876, A
153	7	3.2	362	4	US-10-489-372-19	Sequence 19, Appl	226	7	3.2	871	6	US-11-097-143-1638	Sequence 1638, Ap
154	7	3.2	366	5	US-10-491-823-203	Sequence 203, App	227	7	3.2	876	5	US-10-830-828-6	Sequence 6, Appl1
155	7	3.2	367	5	US-10-91-823-328	Sequence 328, App	228	7	3.2	876	5	US-10-830-828-6	Sequence 6, Appl1
156	7	3.2	367	5	US-10-732-923-10111	Sequence 10111, A	229	7	3.2	900	4	US-10-408-768A-1502	Sequence 1502, App
157	7	3.2	368	5	US-10-732-923-10107	Sequence 10107, A	230	7	3.2	923	5	US-10-732-923-18453	Sequence 18453, A
158	7	3.2	368	5	US-10-732-923-10108	Sequence 10108, A	231	7	3.2	956	4	US-10-108-260A-3037	Sequence 3037, Ap
159	7	3.2	368	5	US-10-732-923-10109	Sequence 10109, A	232	7	3.2	960	4	US-10-437-963-187517	Sequence 187517,
160	7	3.2	373	4	US-10-701-919-4	Sequence 4, Appl1	233	7	3.2	977	4	US-10-425-115-272448	Sequence 272448,
161	7	3.2	379	4	US-10-425-115-202498	Sequence 202498,	234	7	3.2	1045	6	US-11-097-143-13350	Sequence 13350, A
162	7	3.2	389	4	US-10-437-963-199329	Sequence 199329,	235	7	3.2	1074	4	US-10-389-566-2371	Sequence 2371, App
163	7	3.2	399	4	US-10-369-493-13923	Sequence 13923, A	236	7	3.2	1088	4	US-10-424-599-191750	Sequence 191750,
164	7	3.2	402	4	US-10-437-963-188201	Sequence 188201,	237	7	3.2	1091	3	US-10-282-122A-73976	Sequence 73976, A
165	7	3.2	410	4	US-10-437-963-182626	Sequence 182626,	238	7	3.2	1099	3	US-09-769-787-11	Sequence 11, Appl
166	7	3.2	418	4	US-10-424-599-240813	Sequence 240813,	239	7	3.2	1100	5	US-10-472-928-2300	Sequence 2300, Ap
167	7	3.2	422	4	US-10-128-714-3539	Sequence 3539, Ap	240	7	3.2	1100	5	US-10-494-940-22	Sequence 22, Appl
168	7	3.2	422	4	US-10-128-714-8539	Sequence 8539, Ap	241	7	3.2	1102	5	US-10-617-320-4859	Sequence 4859, Ap
169	7	3.2	426	5	US-10-501-282-5656	Sequence 5656, Ap	242	7	3.2	1118	5	US-10-732-923-8271	Sequence 8271, Ap
170	7	3.2	431	4	US-10-424-599-148023	Sequence 148023,	243	7	3.2	1135	4	US-09-972-221-108	Sequence 108, App
171	7	3.2	433	4	US-10-309-175-4	Sequence 4, Appl1	244	7	3.2	1135	4	US-10-096-628-108	Sequence 108, App
172	7	3.2	433	4	US-10-437-963-118208	Sequence 118208,	245	7	3.2	1146	4	US-10-316-532-2	Sequence 2, Appl1
173	7	3.2	434	4	US-10-424-599-148200	Sequence 148200,	246	7	3.2	1150	4	US-10-316-533-4	Sequence 4, Appl1

247	7	3.2	1150	4	US-10-433-544-4	Sequence 4, Appl1	320	2.7	39	3	US-09-864-761-41697	Sequence 41697, A
248	7	3.2	1153	5	US-10-828-815-133	Sequence 133, Appl	321	2.7	39	4	US-10-312-691-4	Sequence 4, Appl1
249	7	3.2	1161	4	US-10-032-189-133	Sequence 133, App	322	2.7	39	4	US-10-425-115-33435	Sequence 33435, A
250	7	3.2	1162	4	US-10-094-749-2705	Sequence 2705, Ap	323	2.7	40	4	US-10-038-612-56363	Sequence 56, Appl
251	7	3.2	1172	5	US-10-828-815-35	Sequence 35, Appl	324	2.7	40	4	US-10-425-115-185768	Sequence 185768, A
252	7	3.2	1179	3	US-09-972-211-109	Sequence 109, Appl	325	2.7	41	4	US-10-029-386-27473	Sequence 27473, A
253	7	3.2	1179	4	US-10-096-625-109	Sequence 109, App	326	2.7	42	3	US-09-864-761-48522	Sequence 48522, A
254	7	3.2	1202	4	US-10-408-765A-997	Sequence 997, App	327	2.7	42	3	US-09-917-340-69	Sequence 69, Appl
255	7	3.2	1265	4	US-10-263-929-127	Sequence 127, App	328	2.7	42	5	US-10-844-837-69	Sequence 69, Appl
256	7	3.2	1285	4	US-10-205-823-273	Sequence 273, App	329	2.7	42	5	US-10-909-119-83	Sequence 83, Appl
257	7	3.2	1285	4	US-10-177-293-317	Sequence 317, App	330	2.7	42	5	US-10-657-851-69	Sequence 69, Appl
258	7	3.2	1285	4	US-10-263-929-126	Sequence 126, App	331	2.7	44	5	US-10-794-514A-486	Sequence 486, App
259	7	3.2	1285	6	US-11-051-454-273	Sequence 273, App	332	2.7	45	4	US-10-008-524A-103	Sequence 103, App
260	7	3.2	1311	5	US-10-501-282-3598	Sequence 3598, Ap	333	2.7	45	4	US-10-350-719-103	Sequence 103, App
261	7	3.2	1314	5	US-10-756-149-5572	Sequence 5572, Ap	334	2.7	45	4	US-10-425-115-18666	Sequence 18666, A
262	7	3.2	1336	4	US-10-369-493-1723	Sequence 1723, Ap	335	2.7	45	4	US-10-425-115-35529	Sequence 35529, A
263	7	3.2	1383	5	US-10-501-282-3600	Sequence 3600, Ap	336	2.7	46	4	US-10-724-912A-6620	Sequence 6620, Ap
264	7	3.2	1407	5	US-10-732-923-3359	Sequence 3359, Ap	337	2.7	47	4	US-10-437-963-147175	Sequence 147175, A
265	7	3.2	1583	4	US-10-437-963-159486	Sequence 159486, A	338	2.7	48	4	US-10-425-115-334970	Sequence 334970, A
266	7	3.2	1644	4	US-10-408-765A-919	Sequence 919, App	339	2.7	49	4	US-10-424-599-253504	Sequence 253504, A
267	7	3.2	2285	5	US-10-773-446-101	Sequence 101, App	340	2.7	50	4	US-10-425-115-218210	Sequence 218210, A
268	7	3.2	2429	5	US-10-732-923-18427	Sequence 18427, A	341	2.7	51	3	US-09-864-761-35058	Sequence 35058, A
269	7	3.2	2440	5	US-10-732-923-18452	Sequence 18452, A	342	2.7	52	3	US-09-971-980-57	Sequence 57, Appl
270	7	3.2	2441	4	US-10-109-886-8	Sequence 8, Appl1	343	2.7	52	4	US-10-437-963-134817	Sequence 134817, A
271	7	3.2	2441	4	US-10-628-957-2	Sequence 2, Appl1	344	2.7	52	4	US-10-437-963-193623	Sequence 193623, A
272	7	3.2	2441	5	US-10-473-127-643	Sequence 643, App	345	2.7	52	5	US-10-966-576-67	Sequence 67, Appl
273	7	3.2	2441	5	US-10-732-923-18428	Sequence 18428, A	346	2.7	53	4	US-10-028-386-29940	Sequence 29940, A
274	7	3.2	2441	5	US-10-732-923-18429	Sequence 18429, A	347	2.7	53	4	US-10-425-115-296334	Sequence 296334, A
275	7	3.2	2442	4	US-10-109-886-10	Sequence 10, Appl	348	2.7	54	4	US-10-424-599-270787	Sequence 270787, A
276	7	3.2	2442	4	US-10-473-127-631	Sequence 631, App	349	2.7	54	4	US-10-425-115-286486	Sequence 286486, A
277	7	3.2	2442	5	US-10-473-127-633	Sequence 633, App	350	2.7	54	4	US-10-425-115-162961	Sequence 329691, A
278	7	3.2	2442	5	US-10-473-127-645	Sequence 645, App	351	2.7	55	4	US-10-424-599-204650	Sequence 204650, A
279	7	3.2	2442	5	US-10-732-923-18450	Sequence 647, App	352	2.7	55	4	US-10-425-115-271198	Sequence 271198, A
280	7	3.2	2442	5	US-10-732-923-18450	Sequence 18450, A	353	2.7	56	3	US-09-864-761-34621	Sequence 34621, A
281	7	3.2	2442	5	US-10-732-923-18451	Sequence 18451, A	354	2.7	58	3	US-09-971-980-69	Sequence 69, Appl
282	7	3.2	2442	5	US-10-840-060-92	Sequence 92, Appl	355	2.7	58	3	US-09-945-917-38	Sequence 38, Appl
283	7	3.2	2442	5	US-10-200-708-103	Sequence 103, App	356	2.7	58	4	US-10-424-599-143876	Sequence 143876, A
284	7	2.7	10	4	US-10-200-708-128	Sequence 128, App	357	2.7	58	5	US-10-966-576-69	Sequence 69, Appl
285	6	2.7	10	4	US-10-200-708-209	Sequence 209, App	358	2.7	59	3	US-09-738-626-6552	Sequence 6552, Ap
286	6	2.7	10	4	US-10-200-708-400	Sequence 400, App	359	2.7	59	4	US-10-437-963-129564	Sequence 129564, A
287	6	2.7	15	3	US-09-876-904A-291	Sequence 291, App	360	2.7	59	5	US-10-450-763-34284	Sequence 34284, A
288	6	2.7	15	3	US-09-461-093-23	Sequence 23, Appl	361	2.7	60	4	US-10-424-599-229175	Sequence 229175, A
289	6	2.7	18	5	US-09-315-355-6	Sequence 6, Appl1	362	2.7	60	4	US-10-425-115-190040	Sequence 190040, A
290	6	2.7	18	5	US-10-848-572-6	Sequence 6, Appl1	363	2.7	60	5	US-10-450-763-11272	Sequence 11272, A
291	6	2.7	19	3	US-09-864-761-40260	Sequence 40260, A	364	2.7	62	4	US-10-767-701-52893	Sequence 52893, A
292	6	2.7	19	3	US-10-813-977-5	Sequence 5, Appl1	365	2.7	62	4	US-10-425-115-334122	Sequence 334122, A
293	6	2.7	20	5	US-09-864-761-44034	Sequence 44034, A	366	2.7	62	4	US-10-425-115-333817	Sequence 333817, A
294	6	2.7	20	5	US-10-478-112A-2	Sequence 2, Appl1	367	2.7	65	4	US-10-424-599-161213	Sequence 161213, A
295	6	2.7	20	5	US-10-478-112A-3	Sequence 3, Appl1	368	2.7	65	4	US-10-437-963-133003	Sequence 133003, A
296	6	2.7	20	5	US-10-478-112A-4	Sequence 4, Appl1	369	2.7	65	4	US-10-425-115-241014	Sequence 241014, A
297	6	2.7	20	5	US-10-478-112A-8	Sequence 8, Appl1	370	2.7	65	5	US-10-450-763-37255	Sequence 37255, A
298	6	2.7	20	5	US-10-478-112A-9	Sequence 9, Appl1	371	2.7	66	3	US-09-989-920-219	Sequence 219, App
299	6	2.7	20	5	US-10-478-112A-10	Sequence 10, Appl	372	2.7	67	3	US-09-864-761-33506	Sequence 33506, A
300	6	2.7	20	5	US-10-862-195-677	Sequence 677, App	373	2.7	67	3	US-09-864-761-40606	Sequence 40606, A
301	6	2.7	20	5	US-10-862-195-678	Sequence 678, App	374	2.7	67	4	US-10-425-115-260163	Sequence 260163, A
302	6	2.7	21	3	US-09-962-756-48	Sequence 48, App	375	2.7	67	5	US-10-732-923-31747	Sequence 31747, A
303	6	2.7	21	4	US-10-253-471-48	Sequence 48, Appl	376	2.7	67	6	US-11-037-143-32763	Sequence 32763, A
304	6	2.7	21	4	US-10-253-493-48	Sequence 48, Appl	377	2.7	68	3	US-09-971-980-63	Sequence 63, Appl
305	6	2.7	21	4	US-09-983-802-305	Sequence 305, App	378	2.7	68	4	US-10-424-599-166978	Sequence 166978, A
306	6	2.7	23	3	US-09-984-490-305	Sequence 305, App	379	2.7	68	4	US-10-424-599-206476	Sequence 206476, A
307	6	2.7	23	3	US-09-973-278-483	Sequence 483, App	380	2.7	68	4	US-10-424-599-274097	Sequence 274097, A
308	6	2.7	25	3	US-09-971-980-71	Sequence 71, Appl	381	2.7	68	4	US-10-425-115-191482	Sequence 191482, A
309	6	2.7	25	4	US-10-053-485-45	Sequence 45, Appl	382	2.7	68	5	US-10-966-576-63	Sequence 63, Appl
310	6	2.7	25	4	US-10-966-576-71	Sequence 71, Appl	383	2.7	69	4	US-10-767-701-454517	Sequence 454517, A
311	6	2.7	30	6	US-11-066-697-1238	Sequence 1238, Ap	384	2.7	69	4	US-10-767-701-60703	Sequence 60703, A
312	6	2.7	31	3	US-09-971-980-61	Sequence 61, Appl	385	2.7	69	4	US-10-425-115-261687	Sequence 261687, A
313	6	2.7	31	5	US-10-966-576-61	Sequence 61, Appl	386	2.7	70	4	US-10-106-686-5831	Sequence 5831, Ap
314	6	2.7	33	4	US-10-425-115-234358	Sequence 234358, A	387	2.7	70	4	US-10-029-386-27874	Sequence 27874, A
315	6	2.7	34	5	US-10-450-763-32778	Sequence 32778, A	388	2.7	70	4	US-10-424-599-229358	Sequence 229358, A
316	6	2.7	35	4	US-10-339-740-254	Sequence 254, App	389	2.7	70	4	US-10-425-115-207725	Sequence 307725, A
317	6	2.7	35	4	US-10-425-115-204044	Sequence 204044, A	390	2.7	70	4	US-10-425-115-440163	Sequence 340163, A
318	6	2.7	36	4	US-10-351-641-760	Sequence 760, App	391	2.7	71	4	US-10-424-599-183419	Sequence 183419, A
319	6	2.7	37	4	US-10-425-115-355461	Sequence 355461, A	392	2.7	72	4	US-10-424-599-180657	Sequence 180657, A

333	6	2.7	72	4	US-10-424-599-220736	Sequence 220736,	466	6	2.7	89	4	US-10-425-115-338885	Sequence 338885,
334	6	2.7	72	4	US-10-425-115-354682	Sequence 354682,	467	6	2.7	89	4	US-10-425-115-339977	Sequence 339977,
335	6	2.7	73	4	US-10-425-115-185766	Sequence 185766,	468	6	2.7	89	4	US-10-425-115-351281	Sequence 351281,
336	6	2.7	73	4	US-10-425-115-267758	Sequence 267758,	469	6	2.7	89	6	US-11-097-143-32544	Sequence 32544, A
337	6	2.7	74	4	US-10-424-599-271433	Sequence 271433,	470	6	2.7	90	4	US-10-424-599-187302	Sequence 187302,
338	6	2.7	74	4	US-10-424-599-271433	Sequence 271433,	471	6	2.7	90	4	US-10-425-115-197959	Sequence 197959,
339	6	2.7	74	4	US-10-437-963-116225	Sequence 116225,	472	6	2.7	90	4	US-10-425-115-245039	Sequence 245039,
340	6	2.7	74	6	US-11-097-143-29868	Sequence 29868, A	473	6	2.7	91	3	US-09-864-761-40634	Sequence 40634, A
401	6	2.7	75	4	US-10-864-761-40037	Sequence 40037, A	474	6	2.7	91	4	US-10-339-740-10	Sequence 10, Appl
402	6	2.7	75	3	US-10-425-115-191322	Sequence 191322,	475	6	2.7	92	3	US-09-864-408A-3954	Sequence 3954, Appl
403	6	2.7	75	4	US-10-425-115-300651	Sequence 300651,	476	6	2.7	92	4	US-10-424-599-35487	Sequence 35487,
404	6	2.7	76	4	US-10-424-599-241422	Sequence 241422,	477	6	2.7	92	4	US-10-424-599-35487	Sequence 35487,
405	6	2.7	76	4	US-10-424-599-258567	Sequence 258567,	478	6	2.7	92	5	US-10-450-763-42232	Sequence 42232, A
406	6	2.7	76	4	US-10-437-963-153323	Sequence 153323,	479	6	2.7	92	5	US-10-450-763-47747	Sequence 47747, A
407	6	2.7	76	4	US-10-425-115-248509	Sequence 248509,	480	6	2.7	92	6	US-10-450-763-57932	Sequence 57932, A
408	6	2.7	76	4	US-10-425-115-329383	Sequence 329383,	481	6	2.7	93	4	US-11-135-597-209	Sequence 209, Appl
409	6	2.7	76	4	US-10-425-115-340012	Sequence 340012,	482	6	2.7	93	4	US-10-437-963-155946	Sequence 155946,
410	6	2.7	77	4	US-10-424-599-188090	Sequence 188090,	483	6	2.7	93	4	US-10-767-701-54583	Sequence 54583, A
411	6	2.7	77	4	US-10-425-115-293660	Sequence 293660,	484	6	2.7	93	4	US-10-425-115-214447	Sequence 214447, A
412	6	2.7	78	4	US-10-283-618-15	Sequence 15,	485	6	2.7	93	5	US-10-450-763-33327	Sequence 33327, A
413	6	2.7	78	4	US-10-133-709-8	Sequence 8,	486	6	2.7	93	5	US-10-450-763-43747	Sequence 43747, A
414	6	2.7	78	4	US-10-425-114-58878	Sequence 58878, A	487	6	2.7	94	4	US-10-450-763-56871	Sequence 56871, A
415	6	2.7	78	4	US-10-437-963-104660	Sequence 104660,	488	6	2.7	94	4	US-10-424-599-443626	Sequence 443626,
416	6	2.7	78	4	US-10-423-115-325038	Sequence 325038,	489	6	2.7	94	4	US-10-767-701-57226	Sequence 57226, A
417	6	2.7	78	5	US-10-113-709A-8	Sequence 8, Appl	490	6	2.7	95	4	US-10-437-963-118987	Sequence 118987,
418	6	2.7	78	5	US-10-994-915-8	Sequence 8, Appl	491	6	2.7	95	4	US-10-437-963-129100	Sequence 129100,
419	6	2.7	79	4	US-10-424-599-199208	Sequence 199208,	492	6	2.7	95	4	US-10-425-115-194205	Sequence 194205,
420	6	2.7	79	4	US-10-424-599-271880	Sequence 271880,	493	6	2.7	95	6	US-11-135-597-211	Sequence 211, Appl
421	6	2.7	79	4	US-10-425-115-221167	Sequence 221167,	494	6	2.7	95	6	US-11-135-597-215	Sequence 215, Appl
422	6	2.7	79	4	US-10-425-115-266243	Sequence 266243,	495	6	2.7	95	6	US-09-216-393-166	Sequence 166, Appl
423	6	2.7	80	3	US-09-924-358-38	Sequence 38, Appl	496	6	2.7	96	3	US-09-839-329-1	Sequence 1, Appl
424	6	2.7	80	4	US-10-410-764-38	Sequence 38, Appl	497	6	2.7	96	4	US-10-053-261-212	Sequence 212, Appl
425	6	2.7	80	4	US-10-437-963-108889	Sequence 108889,	498	6	2.7	96	4	US-10-190-435-256	Sequence 256, Appl
426	6	2.7	80	4	US-10-767-701-53598	Sequence 53598, A	499	6	2.7	96	4	US-10-190-435-257	Sequence 257, Appl
427	6	2.7	80	4	US-10-425-115-288543	Sequence 288543,	500	6	2.7	96	4	US-10-190-435-264	Sequence 264, Appl
428	6	2.7	80	4	US-10-425-115-323576	Sequence 323576,	501	6	2.7	96	4	US-10-190-435-265	Sequence 265, Appl
429	6	2.7	80	4	US-10-425-115-357682	Sequence 357682,	502	6	2.7	96	4	US-10-190-435-266	Sequence 266, Appl
430	6	2.7	81	3	US-09-764-847-569	Sequence 569, Appl	503	6	2.7	96	4	US-10-190-435-267	Sequence 267, Appl
431	6	2.7	81	3	US-09-864-408A-1470	Sequence 1470, Appl	504	6	2.7	96	4	US-10-190-435-268	Sequence 268, Appl
432	6	2.7	81	4	US-10-092-154-569	Sequence 569, Appl	505	6	2.7	96	4	US-10-190-435-269	Sequence 269, Appl
433	6	2.7	81	4	US-10-437-963-193448	Sequence 193448,	506	6	2.7	96	4	US-10-190-435-270	Sequence 270, Appl
434	6	2.7	81	4	US-10-425-115-358925	Sequence 358925,	507	6	2.7	96	4	US-10-190-435-280	Sequence 280, Appl
435	6	2.7	82	4	US-10-424-599-256315	Sequence 256315,	508	6	2.7	96	4	US-10-190-435-282	Sequence 282, Appl
436	6	2.7	82	4	US-10-425-115-188536	Sequence 188536,	509	6	2.7	96	4	US-10-190-435-284	Sequence 284, Appl
437	6	2.7	82	4	US-10-425-115-197867	Sequence 197867,	510	6	2.7	96	4	US-10-321-856-266	Sequence 266, Appl
438	6	2.7	82	4	US-10-425-115-216185	Sequence 216185,	511	6	2.7	96	4	US-10-424-599-14753	Sequence 14753,
439	6	2.7	82	4	US-10-425-115-293873	Sequence 293873,	512	6	2.7	96	4	US-10-424-599-167620	Sequence 167620,
440	6	2.7	83	4	US-10-264-049-3394	Sequence 3394, Appl	513	6	2.7	96	4	US-10-325-468-36	Sequence 36, Appl
441	6	2.7	83	4	US-10-424-599-267146	Sequence 267146,	514	6	2.7	96	4	US-10-425-115-201684	Sequence 201684,
442	6	2.7	83	5	US-10-437-963-161466	Sequence 161466,	515	6	2.7	96	5	US-10-627-649-272	Sequence 272, Appl
443	6	2.7	83	5	US-10-719-993-573	Sequence 573, Appl	516	6	2.7	96	5	US-10-780-507-65	Sequence 65, Appl
444	6	2.7	84	2	US-08-860-844-77	Sequence 77, Appl	517	6	2.7	96	5	US-10-780-507-66	Sequence 66, Appl
445	6	2.7	84	4	US-10-407-543-77	Sequence 77, Appl	518	6	2.7	96	5	US-10-780-507-67	Sequence 67, Appl
446	6	2.7	84	4	US-10-424-599-166438	Sequence 166438,	519	6	2.7	96	5	US-10-780-507-115	Sequence 115, Appl
447	6	2.7	84	4	US-10-424-599-229022	Sequence 229022,	520	6	2.7	96	5	US-10-780-507-117	Sequence 117, Appl
448	6	2.7	84	4	US-10-425-115-191511	Sequence 191511,	521	6	2.7	96	5	US-11-135-597-206	Sequence 206, Appl
449	6	2.7	84	4	US-10-425-115-285543	Sequence 285543,	522	6	2.7	96	6	US-11-135-597-207	Sequence 207, Appl
450	6	2.7	85	3	US-09-833-245-1385	Sequence 1385, Appl	523	6	2.7	96	6	US-11-135-597-208	Sequence 208, Appl
451	6	2.7	85	3	US-09-833-245-1387	Sequence 1387, Appl	524	6	2.7	96	6	US-11-135-597-210	Sequence 210, Appl
452	6	2.7	85	4	US-10-424-599-157592	Sequence 157592,	525	6	2.7	97	6	US-10-437-963-170229	Sequence 170229,
453	6	2.7	85	4	US-10-437-963-180511	Sequence 180511,	526	6	2.7	97	6	US-10-425-115-259822	Sequence 259822,
454	6	2.7	85	5	US-10-926-683-880	Sequence 880, Appl	527	6	2.7	97	4	US-10-425-115-339281	Sequence 339281,
455	6	2.7	85	5	US-10-450-763-35710	Sequence 35710, A	528	6	2.7	98	4	US-10-425-115-50809	Sequence 50809, A
456	6	2.7	86	4	US-10-424-599-149943	Sequence 149943,	529	6	2.7	98	4	US-10-450-763-40994	Sequence 40994, A
457	6	2.7	87	4	US-10-425-115-190448	Sequence 190448,	530	6	2.7	98	5	US-10-437-963-124130	Sequence 124130,
458	6	2.7	87	6	US-11-073-560-39	Sequence 39, Appl	531	6	2.7	99	4	US-10-320-788-116	Sequence 116, Appl
459	6	2.7	88	3	US-09-764-891-4126	Sequence 4126, Appl	532	6	2.7	100	4	US-10-424-599-168809	Sequence 168809,
460	6	2.7	88	4	US-10-289-762-6848	Sequence 6848, Appl	533	6	2.7	100	4	US-10-425-115-230005	Sequence 230005,
461	6	2.7	88	4	US-10-424-599-158350	Sequence 158350,	534	6	2.7	100	4	US-10-425-115-297520	Sequence 297520,
462	6	2.7	88	4	US-10-424-599-221166	Sequence 221166,	535	6	2.7	101	4	US-10-691-599-2	Sequence 2, Appl
463	6	2.7	88	4	US-10-425-115-262777	Sequence 262777,	536	6	2.7	101	4	US-10-425-115-241382	Sequence 241382,
464	6	2.7	88	4	US-10-425-115-280439	Sequence 280439,	537	6	2.7	101	4	US-10-450-763-54500	Sequence 54500, A
465	6	2.7	89	4	US-10-767-701-62131	Sequence 62131, A	538	6	2.7	101	5		

539	6	2.7	102	3	US-09-833-245-915	Sequence 915, App	612	2.7	113	4	US-10-282-122A-70489	Sequence 70489, A
540	6	2.7	102	4	US-10-437-963-196785	Sequence 196785, A	613	2.7	113	4	US-10-282-122A-71492	Sequence 71492, A
541	6	2.7	102	5	US-10-450-763-35729	Sequence 35729, A	614	2.7	113	4	US-10-424-599-248203	Sequence 248203, A
542	6	2.7	103	4	US-10-424-599-217400	Sequence 217400, A	615	2.7	113	5	US-10-789-102-193	Sequence 193, App
543	6	2.7	103	4	US-10-425-114-50527	Sequence 50527, A	616	2.7	113	5	US-10-864-252-498	Sequence 498, App
544	6	2.7	103	4	US-10-437-963-111486	Sequence 111486, A	617	2.7	113	6	US-11-021-949-191	Sequence 191, App
545	6	2.7	103	4	US-10-437-963-141033	Sequence 141033, A	618	2.7	113	6	US-11-008-889-65	Sequence 65, App1
546	6	2.7	103	4	US-10-437-963-151948	Sequence 151948, A	619	2.7	113	6	US-11-021-928A-88	Sequence 88, App1
547	6	2.7	103	4	US-10-767-701-34416	Sequence 34416, A	620	2.7	114	4	US-10-424-599-221648	Sequence 221648, A
548	6	2.7	103	4	US-10-425-115-189892	Sequence 189892, A	621	2.7	114	4	US-10-276-774-2630	Sequence 2630, App
549	6	2.7	103	4	US-10-425-115-304753	Sequence 304753, A	622	2.7	114	4	US-10-425-115-229382	Sequence 229382, A
550	6	2.7	103	5	US-10-732-923-4655	Sequence 4655, App	623	2.7	115	4	US-10-424-599-221109	Sequence 221109, A
551	6	2.7	104	3	US-09-864-877-1863	Sequence 45171, A	624	2.7	115	4	US-10-424-599-275006	Sequence 275006, A
552	6	2.7	105	3	US-10-764-877-1863	Sequence 1863, App	625	2.7	115	4	US-10-425-115-340498	Sequence 340498, A
553	6	2.7	105	4	US-10-242-515-1863	Sequence 1863, App	626	2.7	115	4	US-10-108-915-16	Sequence 36, App1
554	6	2.7	105	4	US-10-767-701-31876	Sequence 31876, App	627	2.7	116	4	US-10-425-114-47623	Sequence 47623, A
555	6	2.7	105	4	US-10-425-115-204240	Sequence 204240, A	628	2.7	116	4	US-10-437-963-119371	Sequence 119371, A
556	6	2.7	106	4	US-10-424-599-164658	Sequence 164658, A	629	2.7	116	4	US-10-437-963-167856	Sequence 167856, A
557	6	2.7	106	4	US-10-424-599-221358	Sequence 221358, A	630	2.7	116	4	US-10-425-115-168351	Sequence 168351, A
558	6	2.7	106	4	US-10-437-963-147879	Sequence 147879, A	631	2.7	116	5	US-10-450-763-60319	Sequence 60319, A
559	6	2.7	106	4	US-10-437-963-183171	Sequence 183171, A	632	2.7	116	6	US-11-032-643-36	Sequence 36, App1
560	6	2.7	106	4	US-10-767-701-42144	Sequence 42144, A	633	2.7	116	6	US-11-032-400-36	Sequence 36, App1
561	6	2.7	106	4	US-10-425-115-242630	Sequence 242630, A	634	2.7	117	3	US-09-815-242-12433	Sequence 12433, App
562	6	2.7	106	4	US-10-425-115-312419	Sequence 312419, A	635	2.7	117	3	US-09-815-242-12867	Sequence 12867, A
563	6	2.7	106	5	US-10-732-923-4653	Sequence 4653, App	636	2.7	117	3	US-09-815-242-13094	Sequence 13094, A
564	6	2.7	106	5	US-10-450-763-31336	Sequence 31336, App	637	2.7	117	4	US-10-424-599-165340	Sequence 165340, A
565	6	2.7	107	2	US-08-848-303-11	Sequence 11, App1	638	2.7	117	4	US-10-424-599-279162	Sequence 279162, A
566	6	2.7	107	4	US-10-424-599-186922	Sequence 186922, A	639	2.7	117	4	US-10-425-115-227689	Sequence 227689, A
567	6	2.7	107	4	US-10-655-136-11	Sequence 11, App1	640	2.7	117	4	US-10-425-115-235732	Sequence 235732, A
568	6	2.7	107	4	US-10-425-115-324495	Sequence 324495, A	641	2.7	117	4	US-10-425-115-73631	Sequence 73631, A
569	6	2.7	107	5	US-10-947-979-66	Sequence 66, App1	642	2.7	117	4	US-10-425-115-335927	Sequence 335927, A
570	6	2.7	107	5	US-10-947-979-66	Sequence 4650, App	643	2.7	118	4	US-10-106-698-4923	Sequence 4923, App
571	6	2.7	107	5	US-10-732-923-4650	Sequence 4651, App	644	2.7	118	4	US-10-424-599-245516	Sequence 245516, A
572	6	2.7	107	5	US-10-732-923-4651	Sequence 4656, App	645	2.7	118	4	US-10-437-963-121091	Sequence 121091, A
573	6	2.7	107	5	US-10-732-923-4654	Sequence 4656, App	646	2.7	118	4	US-10-425-115-195499	Sequence 195499, A
574	6	2.7	107	5	US-10-732-923-4656	Sequence 4657, App	647	2.7	119	3	US-09-205-658-44	Sequence 44, App
575	6	2.7	107	5	US-10-732-923-4657	Sequence 167, App	648	2.7	119	3	US-09-963-693-244	Sequence 244, App
576	6	2.7	107	6	US-11-021-949-167	Sequence 64, App1	649	2.7	119	3	US-09-864-408A-8630	Sequence 8630, App
577	6	2.7	107	6	US-11-021-928A-64	Sequence 64, App1	650	2.7	119	4	US-10-437-963-129635	Sequence 129635, A
578	6	2.7	108	4	US-10-424-599-165133	Sequence 165133, A	651	2.7	119	4	US-10-437-963-140565	Sequence 140565, A
579	6	2.7	108	4	US-10-437-963-109555	Sequence 109555, A	652	2.7	119	4	US-10-425-115-181023	Sequence 181023, A
580	6	2.7	108	4	US-10-437-963-198750	Sequence 198750, A	653	2.7	119	5	US-10-739-930-8482	Sequence 8482, App
581	6	2.7	108	4	US-10-767-701-53950	Sequence 53950, A	654	2.7	120	3	US-09-767-041-51	Sequence 51, App1
582	6	2.7	108	4	US-10-425-115-278507	Sequence 278507, A	655	2.7	120	4	US-10-424-599-203141	Sequence 203141, A
583	6	2.7	108	5	US-10-947-979-50	Sequence 50, App1	656	2.7	120	4	US-10-424-599-205183	Sequence 205183, A
584	6	2.7	109	4	US-10-630-590-59	Sequence 59, App1	657	2.7	120	4	US-10-437-963-143281	Sequence 143281, A
585	6	2.7	109	4	US-10-424-599-144336	Sequence 144336, A	658	2.7	120	4	US-10-425-115-356991	Sequence 356991, A
586	6	2.7	109	4	US-10-424-599-196176	Sequence 196176, A	659	2.7	121	4	US-10-424-599-218441	Sequence 218441, A
587	6	2.7	109	4	US-10-437-963-117011	Sequence 117011, A	660	2.7	121	4	US-10-437-963-195123	Sequence 195123, A
588	6	2.7	109	4	US-10-437-963-139183	Sequence 139183, A	661	2.7	122	4	US-10-425-115-226336	Sequence 226336, A
589	6	2.7	109	4	US-10-437-963-166711	Sequence 166711, A	662	2.7	123	4	US-10-424-599-154335	Sequence 154335, A
590	6	2.7	109	4	US-10-425-115-345897	Sequence 345897, A	663	2.7	123	4	US-10-276-774-1524	Sequence 1524, App
591	6	2.7	109	5	US-10-789-102-158	Sequence 158, App	664	2.7	123	4	US-10-437-963-170419	Sequence 170419, A
592	6	2.7	109	5	US-10-732-923-4652	Sequence 4652, App	665	2.7	123	4	US-10-767-701-43301	Sequence 43301, A
593	6	2.7	109	5	US-10-450-763-36874	Sequence 36874, A	666	2.7	123	4	US-10-425-115-204881	Sequence 204881, A
594	6	2.7	109	5	US-10-450-763-37317	Sequence 37317, A	667	2.7	123	4	US-10-425-115-26518	Sequence 26518, A
595	6	2.7	110	4	US-10-767-701-52346	Sequence 52346, A	668	2.7	123	4	US-10-424-599-178681	Sequence 178681, A
596	6	2.7	110	4	US-10-425-115-295180	Sequence 295180, A	669	2.7	124	4	US-10-425-114-52462	Sequence 52462, A
597	6	2.7	110	5	US-10-926-543-88	Sequence 88, App1	670	2.7	124	4	US-10-437-963-161960	Sequence 161960, A
598	6	2.7	111	4	US-10-437-963-108545	Sequence 108545, A	671	2.7	124	4	US-10-724-972A-5348	Sequence 5348, App
599	6	2.7	111	4	US-10-437-963-103228	Sequence 103228, A	672	2.7	125	3	US-09-864-761-16218	Sequence 761, App
600	6	2.7	111	4	US-10-437-963-125600	Sequence 125600, A	673	2.7	125	4	US-10-437-963-147613	Sequence 147613, A
601	6	2.7	111	4	US-10-767-701-52346	Sequence 52346, A	674	2.7	125	4	US-10-767-701-52718	Sequence 52718, A
602	6	2.7	111	4	US-10-425-115-248201	Sequence 248201, A	675	2.7	125	4	US-10-425-115-184354	Sequence 184354, A
603	6	2.7	111	4	US-10-425-115-265004	Sequence 265004, A	676	2.7	126	4	US-10-424-599-182253	Sequence 182253, A
604	6	2.7	111	4	US-10-425-115-314574	Sequence 314574, A	677	2.7	126	4	US-10-437-963-161854	Sequence 161854, A
605	6	2.7	112	4	US-10-424-599-155874	Sequence 155874, A	678	2.7	126	4	US-10-437-963-195320	Sequence 195320, A
606	6	2.7	112	4	US-10-800-161-23	Sequence 23, App1	679	2.7	126	4	US-10-437-963-195013	Sequence 195013, A
607	6	2.7	112	5	US-10-760-752-23	Sequence 23, App1	680	2.7	126	4	US-10-437-963-202043	Sequence 202043, A
608	6	2.7	113	4	US-10-101-464A-498	Sequence 498, App	681	2.7	126	5	US-10-384-107-59	Sequence 59, App1
609	6	2.7	113	4	US-10-192-381-65	Sequence 94, App1	682	2.7	127	4	US-10-389-647-592	Sequence 592, App
610	6	2.7	113	4	US-10-630-590-94	Sequence 94, App1	683	2.7	127	4	US-10-425-114-58027	Sequence 58027, A
611	6	2.7	113	4	US-10-282-132A-43939	Sequence 43939, A	684	2.7	127	4	US-10-437-963-151160	Sequence 151160, A

685	6	2.7	127	4	US-10-425-115-208648	Sequence 208648,
686	6	2.7	127	4	US-10-425-115-273999	Sequence 273999,
687	6	2.7	127	5	US-10-450-763-51129	Sequence 51129, A
688	6	2.7	128	3	US-09-933-767-254	Sequence 254, App
689	6	2.7	128	4	US-10-108-605-111	Sequence 111, App
690	6	2.7	128	4	US-10-004-860-254	Sequence 254, App
691	6	2.7	128	4	US-10-023-282-254	Sequence 254, App
692	6	2.7	128	4	US-10-437-963-16526	Sequence 176526,
693	6	2.7	128	4	US-10-767-701-31679	Sequence 31679, A
694	6	2.7	129	4	US-10-106-698-6698	Sequence 6698, App
695	6	2.7	129	4	US-10-424-599-186491	Sequence 186491,
696	6	2.7	129	4	US-10-424-599-257546	Sequence 257546,
697	6	2.7	129	4	US-10-425-114-59315	Sequence 59315, A
698	6	2.7	129	4	US-10-437-963-106103	Sequence 106103,
699	6	2.7	129	4	US-10-437-963-165340	Sequence 165340,
700	6	2.7	130	4	US-10-424-599-100816	Sequence 190816,
701	6	2.7	130	4	US-10-437-963-110074	Sequence 110074,
702	6	2.7	130	4	US-10-437-963-110074	Sequence 136824,
703	6	2.7	130	5	US-10-644-765-198	Sequence 198, App
704	6	2.7	130	5	US-10-644-765-277	Sequence 277, App
705	6	2.7	131	5	US-10-739-930-6037	Sequence 6037, App
706	6	2.7	132	4	US-10-335-977-8933	Sequence 8933, App
707	6	2.7	132	4	US-10-276-774-1373	Sequence 1373, App
708	6	2.7	132	4	US-10-437-963-155496	Sequence 155496,
709	6	2.7	132	4	US-10-425-115-299928	Sequence 299928,
710	6	2.7	133	4	US-10-424-599-263321	Sequence 263321,
711	6	2.7	133	4	US-10-335-977-7691	Sequence 7691, App
712	6	2.7	134	4	US-10-424-599-228191	Sequence 228191,
713	6	2.7	134	4	US-10-425-114-39501	Sequence 39501, A
714	6	2.7	134	4	US-10-437-963-158022	Sequence 168022,
715	6	2.7	134	4	US-10-724-972A-7376	Sequence 7376, App
716	6	2.7	134	4	US-10-425-115-247329	Sequence 247329,
717	6	2.7	135	3	US-09-864-408A-3900	Sequence 3900, App
718	6	2.7	135	4	US-10-264-237-1434	Sequence 1434, App
719	6	2.7	135	4	US-10-425-114-58363	Sequence 58363, A
720	6	2.7	135	4	US-10-425-114-62452	Sequence 62452, A
721	6	2.7	135	4	US-10-425-115-265298	Sequence 265298,
722	6	2.7	136	4	US-10-424-599-144814	Sequence 144814,
723	6	2.7	136	4	US-10-424-599-148075	Sequence 148075,
724	6	2.7	136	4	US-10-424-599-164747	Sequence 164747,
725	6	2.7	136	4	US-10-425-114-56785	Sequence 56785, A
726	6	2.7	136	4	US-10-437-963-143900	Sequence 143900,
727	6	2.7	136	4	US-10-437-963-154137	Sequence 154137,
728	6	2.7	136	4	US-10-767-701-34115	Sequence 34115, A
729	6	2.7	136	4	US-10-767-701-49818	Sequence 49818, A
730	6	2.7	137	4	US-10-424-599-157593	Sequence 157593,
731	6	2.7	137	4	US-10-425-114-67405	Sequence 67405, A
732	6	2.7	137	4	US-10-425-115-24287	Sequence 24287,
733	6	2.7	137	4	US-10-425-115-256701	Sequence 256701,
734	6	2.7	138	4	US-10-424-599-266561	Sequence 266561,
735	6	2.7	138	4	US-10-767-701-33964	Sequence 33964, A
736	6	2.7	138	4	US-10-425-115-196227	Sequence 196227,
737	6	2.7	138	4	US-10-425-115-340207	Sequence 340207,
738	6	2.7	138	5	US-10-424-599-244870	Sequence 244870,
739	6	2.7	139	4	US-10-424-599-141574	Sequence 141574, A
740	6	2.7	139	5	US-10-450-763-40490	Sequence 40297, A
741	6	2.7	139	5	US-10-450-763-40490	Sequence 40490, A
742	6	2.7	140	4	US-10-425-114-40397	Sequence 188897,
743	6	2.7	140	4	US-10-425-115-188997	Sequence 1046, App
744	6	2.7	141	3	US-09-764-868-1046	Sequence 110, App
745	6	2.7	141	4	US-10-076-747-110	Sequence 3194, App
746	6	2.7	141	4	US-10-094-749-31094	Sequence 32045, A
747	6	2.7	141	4	US-10-767-701-32045	Sequence 32492,
748	6	2.7	141	4	US-10-425-115-324972	Sequence 110, App
749	6	2.7	141	6	US-11-057-516-110	Sequence 8010, App
750	6	2.7	142	3	US-09-864-408A-8010	Sequence 933, App
751	6	2.7	142	4	US-10-374-780A-933	Sequence 213560,
752	6	2.7	142	4	US-10-424-599-213560	Sequence 103390,
753	6	2.7	142	4	US-10-437-963-103390	Sequence 53806, A
754	6	2.7	142	4	US-10-767-701-53806	Sequence 36247, A
755	6	2.7	142	5	US-10-450-763-36247	Sequence 32817, A
756	6	2.7	142	6	US-11-097-143-32817	Sequence 61462, A
757	6	2.7	143	4	US-10-425-114-61462	

758	6	2.7	143	4	US-10-767-701-62737	Sequence 62737, A
759	6	2.7	143	6	US-11-040-661-40	Sequence 40, Appl
760	6	2.7	144	4	US-10-101-464A-739	Sequence 739, App
761	6	2.7	144	5	US-10-864-252-739	Sequence 739, App
762	6	2.7	145	4	US-10-156-761-11553	Sequence 11563, A
763	6	2.7	145	4	US-10-424-599-191910	Sequence 191910,
764	6	2.7	145	4	US-10-282-122A-61598	Sequence 61598, A
765	6	2.7	146	4	US-10-424-599-105070	Sequence 210570, A
766	6	2.7	146	4	US-10-767-701-60374	Sequence 60374, A
767	6	2.7	146	5	US-10-450-763-51288	Sequence 51288, A
768	6	2.7	147	4	US-10-437-963-107200	Sequence 107200,
769	6	2.7	147	4	US-10-437-963-180607	Sequence 180607,
770	6	2.7	148	4	US-10-424-599-146677	Sequence 146677,
771	6	2.7	148	4	US-10-425-115-35818	Sequence 92, Appl
772	6	2.7	149	2	US-08-979-847-92	Sequence 8472, App
773	6	2.7	149	4	US-10-156-761-8942	Sequence 92, Appl
774	6	2.7	149	4	US-10-114-104-82	Sequence 92, Appl
775	6	2.7	149	5	US-10-767-701-88700	Sequence 37685, A
776	6	2.7	149	5	US-10-450-763-37685	Sequence 60289, A
777	6	2.7	149	5	US-10-450-763-60289	Sequence 126707,
778	6	2.7	150	4	US-10-437-963-126707	Sequence 146600,
779	6	2.7	150	4	US-10-437-963-155540	Sequence 155540,
780	6	2.7	151	4	US-10-424-599-198157	Sequence 198157,
781	6	2.7	151	4	US-10-437-963-124653	Sequence 124653,
782	6	2.7	151	4	US-10-437-963-126851	Sequence 126851,
783	6	2.7	151	4	US-09-925-300-1007	Sequence 1007, App
784	6	2.7	152	3	US-10-264-049-2211	Sequence 2211, App
785	6	2.7	152	4	US-10-767-701-17025	Sequence 257314,
786	6	2.7	153	4	US-10-424-599-257314	Sequence 36114, A
787	6	2.7	153	4	US-10-767-701-16114	Sequence 22032,
788	6	2.7	154	4	US-10-424-599-22032	Sequence 234583,
789	6	2.7	154	4	US-10-425-115-33453	Sequence 334537,
790	6	2.7	154	4	US-10-425-115-338573	Sequence 338573,
791	6	2.7	154	4	US-10-425-115-339915	Sequence 339915,
792	6	2.7	155	4	US-10-450-763-34458	Sequence 34458, A
793	6	2.7	155	4	US-10-425-114-52390	Sequence 52390, A
794	6	2.7	155	4	US-10-437-963-109555	Sequence 109595,
795	6	2.7	155	4	US-10-437-963-137991	Sequence 174939,
796	6	2.7	155	4	US-10-437-963-174939	Sequence 174939,
797	6	2.7	155	4	US-10-425-115-339915	Sequence 339915,
798	6	2.7	155	6	US-11-097-143-39561	Sequence 39561, A
799	6	2.7	156	4	US-10-369-493-16011	Sequence 16011, A
800	6	2.7	156	4	US-10-424-599-100503	Sequence 180503,
801	6	2.7	156	4	US-10-425-115-285860	Sequence 298860,
802	6	2.7	156	4	US-10-425-115-33327	Sequence 33327,
803	6	2.7	156	4	US-10-450-763-45979	Sequence 45979, A
804	6	2.7	157	4	US-10-437-963-132197	Sequence 132197,
805	6	2.7	158	3	US-09-864-761-38291	Sequence 38291, A
806	6	2.7	158	3	US-09-733-381-22	Sequence 22, Appl
807	6	2.7	158	4	US-10-393-840-46	Sequence 46, Appl
808	6	2.7	158	4	US-10-437-963-109565	Sequence 109565,
809	6	2.7	158	4	US-10-437-963-12082	Sequence 12082,
810	6	2.7	158	4	US-10-437-963-14762	Sequence 164762,
811	6	2.7	158	4	US-10-437-963-185084	Sequence 185084,
812	6	2.7	158	4	US-10-425-115-237506	Sequence 237506,
813	6	2.7	158	4	US-10-425-115-251191	Sequence 251191,
814	6	2.7	158	5	US-10-450-763-47473	Sequence 47473, A
815	6	2.7	158	5	US-11-097-143-10773	Sequence 10773, A
816	6	2.7	159	4	US-10-437-963-151382	Sequence 151382,
817	6	2.7	159	4	US-10-425-115-210034	Sequence 210034,
818	6	2.7	159	5	US-10-450-763-50300	Sequence 50300, A
819	6	2.7	160	4	US-10-425-114-38234	Sequence 38234, A
820	6	2.7	160	4	US-10-425-115-294252	Sequence 294252,
821	6	2.7	161	4	US-10-437-963-111085	Sequence 111085,
822	6	2.7	161	4	US-10-767-701-33149	Sequence 33149, A
823	6	2.7	161	6	US-11-097-143-9315	Sequence 9315, App
824	6	2.7	162	4	US-10-425-114-46333	Sequence 46333, A
825	6	2.7	162	4	US-10-425-114-24109	Sequence 242109,
826	6	2.7	164	4	US-10-425-114-67943	Sequence 67943, A
827	6	2.7	164	4	US-10-767-701-32586	Sequence 32586, A
828	6	2.7	165	4	US-10-767-701-59014	Sequence 59014, A
829	6	2.7	165	4	US-10-425-115-202211	Sequence 202311,
830	6	2.7	166	4	US-10-369-493-15636	Sequence 15636, A

831	6	2.7	166	4	US-10-424-599-258655	Sequence 258655,	904	6	2.7	186	4	US-10-425-115-215429	Sequence 215429,
832	6	2.7	166	4	US-10-437-963-139730	Sequence 139730,	905	6	2.7	187	4	US-10-424-599-156449	Sequence 156449,
833	6	2.7	166	4	US-10-767-701-49182	Sequence 49182, A	906	6	2.7	187	4	US-10-437-963-185697	Sequence 185697,
834	6	2.7	166	4	US-10-425-115-338395	Sequence 338395, A	907	6	2.7	187	4	US-10-767-701-49182	Sequence 49182, A
835	6	2.7	167	4	US-10-156-765-12488	Sequence 12488, A	908	6	2.7	187	5	US-10-450-763-51127	Sequence 51127, A
836	6	2.7	167	4	US-10-245-618-4	Sequence 4, App1	909	6	2.7	187	6	US-11-097-143-543	Sequence 543, App
837	6	2.7	167	4	US-10-437-963-191748	Sequence 191748,	910	6	2.7	188	3	US-09-925-301-1331	Sequence 1331, App
838	6	2.7	168	4	US-10-424-599-259870	Sequence 259870,	911	6	2.7	188	3	US-09-915-582-90	Sequence 90, App1
839	6	2.7	168	4	US-10-767-701-57935	Sequence 57935, A	912	6	2.7	188	4	US-10-277-802-80	Sequence 80, App1
840	6	2.7	168	4	US-10-425-115-212189	Sequence 212189,	913	6	2.7	188	4	US-10-437-963-148630	Sequence 148630,
841	6	2.7	168	4	US-10-425-115-265345	Sequence 265345,	914	6	2.7	188	4	US-10-437-963-178550	Sequence 178550,
842	6	2.7	169	5	US-10-425-115-210032	Sequence 210032,	915	6	2.7	188	5	US-10-896-972-90	Sequence 90, App1
843	6	2.7	169	5	US-10-617-316-261	Sequence 261, App	916	6	2.7	189	4	US-10-450-763-44914	Sequence 44914, A
844	6	2.7	170	4	US-10-282-122A-63981	Sequence 63981, A	917	6	2.7	189	4	US-10-425-115-37210	Sequence 37210, A
845	6	2.7	171	4	US-10-108-260A-4136	Sequence 4136, App	918	6	2.7	189	4	US-10-437-963-177509	Sequence 177509, A
846	6	2.7	171	4	US-10-282-122A-47427	Sequence 47427, A	919	6	2.7	190	4	US-10-767-701-55997	Sequence 55997, A
847	6	2.7	171	4	US-10-425-114-43014	Sequence 43014, A	920	6	2.7	190	4	US-10-425-115-331013	Sequence 331013, A
848	6	2.7	171	4	US-10-437-963-148781	Sequence 148781,	921	6	2.7	190	4	US-10-425-115-337486	Sequence 337486,
849	6	2.7	171	4	US-10-767-701-33797	Sequence 33797, A	922	6	2.7	191	3	US-09-833-245-1875	Sequence 1875, App
850	6	2.7	171	5	US-10-450-763-46203	Sequence 46203, A	923	6	2.7	191	4	US-10-190-435-116	Sequence 116, App
851	6	2.7	172	4	US-10-369-493-19059	Sequence 19059, A	924	6	2.7	191	4	US-10-437-963-135757	Sequence 135757,
852	6	2.7	172	4	US-10-425-115-228000	Sequence 228000,	925	6	2.7	191	4	US-10-437-963-159273	Sequence 159273,
853	6	2.7	172	4	US-10-425-115-350887	Sequence 350887,	926	6	2.7	191	4	US-10-658-834A-875	Sequence 875, App
854	6	2.7	173	4	US-10-461-093-16	Sequence 16, App1	927	6	2.7	191	4	US-10-767-701-31817	Sequence 31817, A
855	6	2.7	173	4	US-10-767-701-49189	Sequence 49189, A	928	6	2.7	191	4	US-10-425-115-168939	Sequence 168939,
856	6	2.7	173	4	US-10-425-115-237504	Sequence 237504,	929	6	2.7	192	4	US-10-425-115-293635	Sequence 293635,
857	6	2.7	173	5	US-10-938-740-13	Sequence 13, App1	930	6	2.7	193	3	US-09-833-245-1873	Sequence 1873, App
858	6	2.7	174	4	US-10-437-963-130412	Sequence 130412,	931	6	2.7	194	4	US-10-437-963-183185	Sequence 183185,
859	6	2.7	175	4	US-10-282-122A-50495	Sequence 50495, A	932	6	2.7	195	4	US-10-424-599-67705	Sequence 67705,
860	6	2.7	175	4	US-10-437-963-168025	Sequence 168025,	933	6	2.7	195	4	US-10-425-115-198731	Sequence 198731,
861	6	2.7	175	4	US-10-425-115-309671	Sequence 309671,	934	6	2.7	195	4	US-10-425-115-154444	Sequence 154444,
862	6	2.7	176	4	US-10-036-729-2	Sequence 2, App1	935	6	2.7	196	3	US-09-815-242-11660	Sequence 11660, A
863	6	2.7	176	4	US-10-461-093-19	Sequence 19, App1	936	6	2.7	196	4	US-10-359-493-17686	Sequence 17686, A
864	6	2.7	176	4	US-10-335-977-7692	Sequence 7692, App	937	6	2.7	196	4	US-10-424-599-100805	Sequence 100805,
865	6	2.7	176	4	US-10-247-843-6	Sequence 6, App1	938	6	2.7	196	4	US-10-425-115-106307	Sequence 106307,
866	6	2.7	176	4	US-10-437-963-122558	Sequence 122558,	939	6	2.7	196	4	US-10-425-115-158639	Sequence 158639,
867	6	2.7	176	4	US-10-425-115-258387	Sequence 258387,	940	6	2.7	196	5	US-10-450-763-165633	Sequence 165633,
868	6	2.7	177	4	US-10-425-115-198734	Sequence 198734,	941	6	2.7	196	5	US-10-450-763-168781	Sequence 168781,
869	6	2.7	178	4	US-10-282-122A-49481	Sequence 49481, A	942	6	2.7	196	5	US-10-450-763-12946	Sequence 12946,
870	6	2.7	178	4	US-10-425-115-229139	Sequence 229139,	943	6	2.7	198	3	US-09-912-020-360	Sequence 360, App
871	6	2.7	179	4	US-10-424-599-241383	Sequence 241383,	944	6	2.7	198	3	US-09-738-626-4812	Sequence 4812, App
872	6	2.7	179	4	US-10-425-114-56318	Sequence 56318, A	945	6	2.7	198	4	US-10-461-093-15	Sequence 15, App1
873	6	2.7	179	4	US-10-425-114-67409	Sequence 67409, A	946	6	2.7	198	5	US-10-771-241-160	Sequence 160, App
874	6	2.7	179	4	US-10-437-963-131063	Sequence 131063,	947	6	2.7	198	5	US-10-489-065A-543	Sequence 543, App
875	6	2.7	180	3	US-09-957-801-6	Sequence 6, App1	948	6	2.7	199	3	US-09-791-497-6	Sequence 6, App1
876	6	2.7	180	4	US-10-282-122A-62334	Sequence 62334, A	949	6	2.7	199	4	US-10-282-122A-46461	Sequence 46461, A
877	6	2.7	180	4	US-10-282-122A-64883	Sequence 64883, A	950	6	2.7	200	3	US-09-957-801-2	Sequence 2, App1
878	6	2.7	180	4	US-10-425-115-202875	Sequence 202875,	951	6	2.7	200	4	US-10-424-599-201127	Sequence 201127,
879	6	2.7	180	4	US-10-425-115-221810	Sequence 221810,	952	6	2.7	200	4	US-10-424-599-229315	Sequence 229315,
880	6	2.7	181	4	US-10-437-963-113816	Sequence 113816,	953	6	2.7	200	4	US-10-437-963-143878	Sequence 143878,
881	6	2.7	181	4	US-10-437-963-120613	Sequence 120613,	954	6	2.7	200	4	US-10-437-963-171377	Sequence 171377,
882	6	2.7	181	4	US-10-425-115-284166	Sequence 284166,	955	6	2.7	200	4	US-10-767-701-15891	Sequence 15891, A
883	6	2.7	183	4	US-10-424-599-263332	Sequence 263332,	956	6	2.7	200	4	US-10-425-115-201901	Sequence 201901,
884	6	2.7	183	4	US-10-425-114-37682	Sequence 37682, A	957	6	2.7	200	4	US-10-017-161-1586	Sequence 1586, App
885	6	2.7	183	4	US-10-437-963-131337	Sequence 131337,	958	6	2.7	201	4	US-10-461-093-18	Sequence 18, App1
886	6	2.7	183	4	US-10-767-701-36803	Sequence 36803, A	959	6	2.7	201	4	US-10-292-798-1262	Sequence 1262, App
887	6	2.7	184	3	US-09-766-826-30	Sequence 30, App1	960	6	2.7	204	4	US-10-424-599-161170	Sequence 161170,
888	6	2.7	184	3	US-09-833-245-347	Sequence 347, App	961	6	2.7	205	4	US-10-425-115-161017	Sequence 161017,
889	6	2.7	184	3	US-09-833-245-2250	Sequence 2250, App	962	6	2.7	205	6	US-11-097-143-15940	Sequence 15940, A
890	6	2.7	184	4	US-10-437-963-189425	Sequence 189425,	963	6	2.7	206	4	US-10-425-114-18316	Sequence 18316, A
891	6	2.7	184	4	US-10-767-701-34765	Sequence 34765, A	964	6	2.7	207	4	US-10-437-963-114172	Sequence 114172,
892	6	2.7	184	4	US-10-425-115-195325	Sequence 195325,	965	6	2.7	208	4	US-10-282-122A-47852	Sequence 47852, A
893	6	2.7	184	4	US-10-425-115-366830	Sequence 366830,	966	6	2.7	208	4	US-10-424-599-166560	Sequence 166560,
894	6	2.7	184	5	US-10-874-484-10	Sequence 10, App1	967	6	2.7	208	4	US-10-437-963-149439	Sequence 149439,
895	6	2.7	185	3	US-09-815-242-11768	Sequence 11768, A	968	6	2.7	208	4	US-10-437-963-190465	Sequence 190465,
896	6	2.7	185	4	US-10-282-122A-66444	Sequence 66444, A	969	6	2.7	208	4	US-10-425-115-213113	Sequence 213113,
897	6	2.7	185	4	US-10-424-599-148731	Sequence 148731,	970	6	2.7	209	4	US-10-425-114-63805	Sequence 63805, A
898	6	2.7	185	4	US-10-335-977-8934	Sequence 8934, App	971	6	2.7	210	4	US-10-437-963-130914	Sequence 130914,
899	6	2.7	185	4	US-10-425-115-257944	Sequence 257944,	972	6	2.7	210	4	US-10-425-115-306129	Sequence 306129,
900	6	2.7	186	4	US-10-424-599-213466	Sequence 213466,	973	6	2.7	210	4	US-10-425-115-313510	Sequence 313510,
901	6	2.7	186	4	US-10-425-114-50370	Sequence 50370, A	974	6	2.7	210	4	US-10-425-115-339994	Sequence 339994,
902	6	2.7	186	4	US-10-437-963-178353	Sequence 178353,	975	6	2.7	211	4	US-10-424-599-233327	Sequence 233327,
903	6	2.7	186	4	US-10-767-701-46427	Sequence 46427, A	976	6	2.7	211	4	US-10-425-114-53589	Sequence 53589, A

```
977 6 2.7 211 4 US-10-767-701-38488 Sequence 38488, A
978 6 2.7 211 6 US-11-097-143-29091 Sequence 29091, A
979 6 2.7 212 4 US-10-767-701-46117 Sequence 46117, A
980 6 2.7 212 6 US-11-097-143-9234 Sequence 9234, Ap
981 6 2.7 213 4 US-10-029-386-32952 Sequence 32952, A
982 6 2.7 213 5 US-10-489-864-71 Sequence 71, Appl
983 6 2.7 213 5 US-10-501-282-1472 Sequence 1472, Ap
984 6 2.7 213 6 US-11-097-143-41157 Sequence 41157, A
985 6 2.7 214 4 US-10-437-963-114044 Sequence 114044, A
986 6 2.7 214 4 US-10-425-115-233480 Sequence 233480, A
987 6 2.7 214 4 US-10-425-115-261828 Sequence 261828, A
988 6 2.7 215 4 US-10-425-114-51176 Sequence 144, Ap
989 6 2.7 215 4 US-10-437-963-138396 Sequence 138396, A
990 6 2.7 215 5 US-10-739-930-7330 Sequence 7330, Ap
991 6 2.7 215 5 US-10-489-864-70 Sequence 70, Appl
992 6 2.7 215 6 US-11-004-418A-15 Sequence 15, Appl
993 6 2.7 216 4 US-10-240-145-144 Sequence 144, Ap
994 6 2.7 216 4 US-10-424-599-239247 Sequence 239247, A
995 6 2.7 216 4 US-10-767-701-44085 Sequence 44085, A
996 6 2.7 216 5 US-10-291-128-144 Sequence 144, Appl
997 6 2.7 216 6 US-10-953-399-42 Sequence 42, Appl
998 6 2.7 217 6 US-11-097-143-36258 Sequence 36258, A
999 6 2.7 217 4 US-10-424-599-194225 Sequence 194225, A
1000 6 2.7 217 4 US-10-276-774-1679 Sequence 1679, Ap
```

ALIGNMENTS

```
RESULT 1
US-10-695-994-7
; Sequence 7, Application US/10695994
; Publication No. US20040142354A1
; GENERAL INFORMATION:
; APPLICANT: MISHRA, LOPA
; TITLE OF INVENTION: GENES CODING FOR EARLY LIVER DEVELOPMENT...
; FILE REFERENCE: P04470051/BAS
; CURRENT FILING DATE: 2003-10-30
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 09/431,184
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US98/08656
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: US 08/841,349
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-695-994-7
```

```
Query Match 100.0%; Score 220; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 7.5e-201;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MELORTSSISGPISPYTGOVPYNNQLBGRFKQLQDEREAVOKKFTKWNSHLARVSC 60
DB 1 MELORTSSISGPISPYTGOVPYNNQLBGRFKQLQDEREAVOKKFTKWNSHLARVSC 60
QY 61 RITDLYTDLRDGMILIKLLEVLSEGERLPPYTKGRMRHICLENVDKALQFLKEGRVHL 120
DB 61 RITDLYTDLRDGMILIKLLEVLSEGERLPPYTKGRMRHICLENVDKALQFLKEGRVHL 120
QY 121 GSHDIYDGNHRLTTLELLEVRROQEBEERKRRPPSPDPTKYSSEAEASQOMDTSKGD 180
DB 121 GSHDIYDGNHRLTTLELLEVRROQEBEERKRRPPSPDPTKYSSEAEASQOMDTSKGD 180
QY 181 QNGLPABQSGPRVSYRSQTYQNKFNRSRTASDHSMG 220
DB 181 QNGLPABQSGPRVSYRSQTYQNKFNRSRTASDHSMG 220
```

```
RESULT 2
US-10-695-994-4
; Sequence 4, Application US/10695994
; Publication No. US20040142354A1
; GENERAL INFORMATION:
; APPLICANT: MISHRA, LOPA
; TITLE OF INVENTION: GENES CODING FOR EARLY LIVER DEVELOPMENT...
; FILE REFERENCE: P04470051/BAS
; CURRENT FILING DATE: 2003-10-30
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 09/431,184
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US98/08656
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: US 08/841,349
; PRIOR FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 2154
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-695-994-4
```

```
Query Match 56.4%; Score 124; DB 4; Length 2154;
Best Local Similarity 100.0%; Pred. No. 1.2e-108;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 10 SGPLSPATYGOVPYNNQLBGRFKQLQDEREAVOKKFTKWNSHLARVSCRITDLYTDL 69
DB 10 SGPLSPATYGOVPYNNQLBGRFKQLQDEREAVOKKFTKWNSHLARVSCRITDLYTDL 69
QY 70 RDGRMLIKLLEVLSEGERLPPYTKGRMRHICLENVDKALQFLKEGRVHLNMGSHDIVDGN 129
DB 70 RDGRMLIKLLEVLSEGERLPPYTKGRMRHICLENVDKALQFLKEGRVHLNMGSHDIVDGN 129
QY 130 HRLT 133
DB 130 HRLT 133
```

```
RESULT 3
US-10-805-684-150
; Sequence 150, Application US/10805684
; Publication No. US20050100966A1
; GENERAL INFORMATION:
; APPLICANT: SAKAMOTO, TAKESHI
; APPLICANT: TAKEDA, SHIZU
; TITLE OF INVENTION: PHOS-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: 58748 (70342)
; CURRENT FILING DATE: 2004-03-19
; PRIOR FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/455,766
; PRIOR FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/459,936
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/460,103
; PRIOR FILING DATE: 2003-04-02
; NUMBER OF SEQ ID NOS: 161
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 150
; LENGTH: 2154
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-805-684-150
```

```
Query Match 56.4%; Score 124; DB 5; Length 2154;
Best Local Similarity 100.0%; Pred. No. 1.2e-108;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 SGPLSPATYGOVPYNNQLBGRFKQLQDEREAVOKKFTKWNSHLARVSCRITDLYTDL 69
```


Db 10 SGPSPAYTGQVYNNYNOEGFQKQLODERBAVOKKPTKWNVSHLAVSCRTITDLYTDL 69
Qy 70 RDGRMLIKLEVLSEGRILPKPTKGMRIHCLENVKALQFLKEQVHLENGSHDIVDGN 129
Db 70 RDGRMLIKLEVLSEGRILPKPTKGMRIHCLENVKALQFLKEQVHLENGSHDIVDGN 129
Qy 130 HRLT 133
Db 130 HRLT 133

RESULT 4
US-10-408-765A-2431

; Sequence 2431, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Boin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2431
; LENGTH: 1630
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2431

Query Match 44.1%; Score 97; DB 4; Length 1630;
Best Local Similarity 100.0%; Pred. No. 4.7e-83;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 DERBAVOKKPTKWNVSHLAVSCRTITDLYTDLRDGRMLIKLEVLSEGRILPKPTKGMRIHCLENVKALQFLKEQVHLENGSHDIVDGNHRLT 133
Db 50 DERBAVOKKPTKWNVSHLAVSCRTITDLYTDLRDGRMLIKLEVLSEGRILPKPTKGMRIHCLENVKALQFLKEQVHLENGSHDIVDGNHRLT 109
Qy 97 IHCLENVDKALQFLKEQVHLENGSHDIVDGNHRLT 133
Db 110 IHCLENVDKALQFLKEQVHLENGSHDIVDGNHRLT 146

RESULT 5
US-10-170-385-15

; Sequence 15, Application US/10170385
; Publication No. US20030203372A1
; GENERAL INFORMATION:
; APPLICANT: Ward, Neil Raymond
; APPLICANT: Mundy, Christopher Robert
; APPLICANT: Kan, On
; APPLICANT: Harris, Robert Alan
; APPLICANT: White, Jonathan
; APPLICANT: Binley, Katie Mary
; APPLICANT: Rayner, William Nigel
; APPLICANT: Naylor, Stuart
; APPLICANT: Kingsman, Susan Mary
; APPLICANT: Krige, David
; TITLE OF INVENTION: ANALYSIS METHOD
; FILE REFERENCE: 532682000100
; CURRENT APPLICATION NUMBER: US/10/170,385
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: PCT/GB02/01662
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: PCT/GB01/05458
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 549

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2364
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-170-385-15

Query Match 44.1%; Score 97; DB 4; Length 2364;
Best Local Similarity 100.0%; Pred. No. 6.5e-83;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 DERBAVOKKPTKWNVSHLAVSCRTITDLYTDLRDGRMLIKLEVLSEGRILPKPTKGMRIHCLENVKALQFLKEQVHLENGSHDIVDGNHRLT 133
Db 50 DERBAVOKKPTKWNVSHLAVSCRTITDLYTDLRDGRMLIKLEVLSEGRILPKPTKGMRIHCLENVKALQFLKEQVHLENGSHDIVDGNHRLT 109
Qy 97 IHCLENVDKALQFLKEQVHLENGSHDIVDGNHRLT 133
Db 110 IHCLENVDKALQFLKEQVHLENGSHDIVDGNHRLT 146

RESULT 6
US-09-945-917-43

; Sequence 43, Application US/09945917
; Publication No. US20030042381A1
; GENERAL INFORMATION:
; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandekerckhove, Joël
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; FILE REFERENCE: P/14-1
; CURRENT APPLICATION NUMBER: US/09/945,917
; CURRENT FILING DATE: 1998-09-21
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-945-917-43

Query Match 17.3%; Score 38; DB 3; Length 57;
Best Local Similarity 100.0%; Pred. No. 3.7e-28;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 MRHICLENVDKALQFLKEQVHLENGSHDIVDGNHRLT 132
Db 18 MRHICLENVDKALQFLKEQVHLENGSHDIVDGNHRLT 55

RESULT 7
US-10-408-765A-2093

; Sequence 2093, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Boin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2093
; LENGTH: 2106
; TYPE: PRT

```
/ ORGANISM: Homo sapiens
US-10-408-765A-2093

Query Match      17.3%; Score 38; DB 4; Length 2106;
Best Local Similarity 100.0%; Pred. No. 8.3e-27;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      95 MRHCLENVDKALQFLKEQRVHLENNGSHDIVDGNHRL 132
Db      108 MRHCLENVDKALQFLKEQRVHLENNGSHDIVDGNHRL 145

RESULT 8
US-10-756-149-4806
; Sequence 4806, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nacasha
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; CURRENT APPLICATION NUMBER: US/10/756,149
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4806
; LENGTH: 2137
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-4806

Query Match      17.3%; Score 38; DB 5; Length 2137;
Best Local Similarity 100.0%; Pred. No. 8.4e-27;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      95 MRHCLENVDKALQFLKEQRVHLENNGSHDIVDGNHRL 132
Db      108 MRHCLENVDKALQFLKEQRVHLENNGSHDIVDGNHRL 145

RESULT 9
US-10-972-637-2
; Sequence 2, Application US/10972637
; Publication No. US20050106659A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: ROTHESTEIN, Jeffrey D.
; APPLICANT: JACKSON, Mandy
; APPLICANT: LIN, Glen
; APPLICANT: LAW, Robert
; APPLICANT: ORLOV, Irina
; TITLE OF INVENTION: GLUTAMATE TRANSPORTER ASSOCIATED PROTEINS AND METHODS OF USE THEREIN
; FILE REFERENCE: JHU650-2
; CURRENT APPLICATION NUMBER: US/10/972,637
; PRIOR FILING DATE: 2004-10-25
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: US 60/161,007
; PRIOR FILING DATE: 1999-10-23
; PRIOR APPLICATION NUMBER: US 60/206,157
; PRIOR FILING DATE: 2000-05-22
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 2388
; TYPE: PRT
; ORGANISM: Rattus
US-10-972-637-2

Query Match      12.3%; Score 27; DB 5; Length 2388;
Best Local Similarity 100.0%; Pred. No. 2.7e-16;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      87 LPKPTGMRHICLENVDKALQFLKEQ 113
Db      103 LPKPTGMRHICLENVDKALQFLKEQ 129

RESULT 10
US-10-756-149-5749
; Sequence 5749, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nacasha
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; FILE REFERENCE: file
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; CURRENT APPLICATION NUMBER: US/10/756,149
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5749
; LENGTH: 2390
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-5749

Query Match      12.3%; Score 27; DB 5; Length 2390;
Best Local Similarity 100.0%; Pred. No. 2.7e-16;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      87 LPKPTGMRHICLENVDKALQFLKEQ 113
Db      103 LPKPTGMRHICLENVDKALQFLKEQ 129

RESULT 11
US-09-945-917-40
; Sequence 40, Application US/09945917
; Publication No. US20030042381A1
; GENERAL INFORMATION:
; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandekerckhove, Joel
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; FILE REFERENCE: P/14-1
; CURRENT APPLICATION NUMBER: US/09/945,917
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-945-917-40

Query Match      11.4%; Score 25; DB 3; Length 51;
Best Local Similarity 100.0%; Pred. No. 8e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      42 VOKTFTKWNVNSHLARVSGRTIDLY 66
Db      16 VOKTFTKWNVNSHLARVSGRTIDLY 40

RESULT 12
US-10-264-049-2841
; Sequence 2841, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Blise et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
FILE REFERENCE: PA133P1
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2841
LENGTH: 806
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (9)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (11)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
FEATURE:
NAME/KEY: MISC_FEATURE
LOCATION: (19)
OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-2841
```

```
Query Match 10.5%; Score 23; DB 4; Length 806;
Best Local Similarity 100.0%; Pred. No. 6.9e-13;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 131 RLTTLELVRRQGEERKRRP 153
Db 699 RLTTLELVRRQGEERKRRP 721
```

```
RESULT 13
US-10-756-149-4730
Sequence 4730, Application US/10756149
Publication No. US20050181375A1
GENERAL INFORMATION:
APPLICANT: Aziz, Natasha
APPLICANT: Zlotnick, Albert
TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
FILE REFERENCE: file
CURRENT APPLICATION NUMBER: US/10/756,149
CURRENT FILING DATE: 2004-01-12
NUMBER OF SEQ ID NOS: 5818
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4730
LENGTH: 2559
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-756-149-4730
```

```
Query Match 10.0%; Score 22; DB 5; Length 2559;
Best Local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 37 DERBAVOKKTKWNSHLARY 58
Db 57 DERBAVOKKTKWNSHLARY 78
```

```
RESULT 14
US-11-097-143-12420
Sequence 12420, Application US/11097143
Publication No. US20050208558A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
```

```
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
TITLE OF INVENTION: DROSOPHILA GENES.
FILE REFERENCE: CLO00728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12420
LENGTH: 2291
TYPE: PRT
ORGANISM: DROSOPHILA
US-11-097-143-12420
```

```
Query Match 8.6%; Score 19; DB 6; Length 2291;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 75 LKLELVISGERLPKPTKG 93
Db 84 LKLELVISGERLPKPTKG 102
```

```
RESULT 15
US-09-945-917-46
Sequence 46, Application US/09945917
Publication No. US20030042381A1
GENERAL INFORMATION:
APPLICANT: Bogaert, Thierry
APPLICANT: Vandekerckhove, Joel
TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
FILE REFERENCE: P/14-1
CURRENT APPLICATION NUMBER: US/09/945,917
CURRENT FILING DATE: 1998-09-21
NUMBER OF SEQ ID NOS: 58
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 48
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-945-917-46
```

```
Query Match 7.3%; Score 16; DB 3; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 107 LQFLKEQRVHLENNGS 122
Db 33 LQFLKEQRVHLENNGS 48
```

```
RESULT 16
US-10-862-195-1550
Sequence 1550, Application US/10862195
Publication No. US20050164324A1
```

GENERAL INFORMATION:
APPLICANT: GIGI, STEVEN P.
TITLE OF INVENTION: SYSTEMS, METHODS AND KITS FOR CHARACTERIZING PHOSPHOPROTEOMES
FILE REFERENCE: 58890(70207)
CURRENT APPLICATION NUMBER: US/10/862,195
CURRENT FILING DATE: 2004-06-04
PRIOR APPLICATION NUMBER: 60/476,010
PRIOR FILING DATE: 2003-06-04
NUMBER OF SEQ ID NOS: 2245
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1550
LENGTH: 17
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: See specification as filed for preferred embodiments
US-10-862-195-1550

Query Match 6.8%; Score 15; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 QVSQNGLPABQGSFR 192
DB 3 QVSQNGLPABQGSFR 17

RESULT 17
US-10-862-195-1551
Sequence 1551, Application US/10862195
Publication No. US20050164324A1
GENERAL INFORMATION:
APPLICANT: GIGI, STEVEN P.
TITLE OF INVENTION: SYSTEMS, METHODS AND KITS FOR CHARACTERIZING PHOSPHOPROTEOMES
FILE REFERENCE: 58890(70207)
CURRENT APPLICATION NUMBER: US/10/862,195
CURRENT FILING DATE: 2004-06-04
PRIOR APPLICATION NUMBER: 60/476,010
PRIOR FILING DATE: 2003-06-04
NUMBER OF SEQ ID NOS: 2245
SOFTWARE: PatentIn version 3.2
SEQ ID NO 1551
LENGTH: 17
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: See specification as filed for preferred embodiments
US-10-862-195-1551

Query Match 6.8%; Score 15; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 QVSQNGLPABQGSFR 192
DB 3 QVSQNGLPABQGSFR 17

RESULT 18
US-10-695-994-20
Sequence 20, Application US/10695994
Publication No. US20040142354A1
GENERAL INFORMATION:
APPLICANT: MISHRA, LOPI
TITLE OF INVENTION: GENES CODING FOR EARLY LIVER DEVELOPMENT...
FILE REFERENCE: P04470US1/BAS
CURRENT APPLICATION NUMBER: US/10/695,994
CURRENT FILING DATE: 2003-10-30
PRIOR APPLICATION NUMBER: 09/431,184
PRIOR FILING DATE: 1999-11-01
PRIOR APPLICATION NUMBER: PCT/US98/08656

PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: US 08/841,349
PRIOR FILING DATE: 1997-04-30
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.0
SEQ ID NO 20
LENGTH: 14
TYPE: PRT
ORGANISM: Mus musculus
US-10-695-994-20

Query Match 6.4%; Score 14; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 7.8e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 FNSRRTASDHSMSG 219
DB 1 FNSRRTASDHSMSG 14

RESULT 19
US-10-369-493-6374
Sequence 6374, Application US/10369493
Publication No. US2003023675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 6374
LENGTH: 2326
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-6374

Query Match 6.4%; Score 14; DB 4; Length 2326;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKTFYKVNSHL 55
DB 110 VOKKTFYKVNSHL 123

RESULT 20
US-10-369-493-5251
Sequence 5251, Application US/10369493
Publication No. US2003023675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 5251
LENGTH: 1355

TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-5251

Query Match
Best Local Similarity 100.0%; Score 11; DB 4; Length 1355;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 VOKKTFTKWN 52
|||||
Db 79 VOKKTFTKWN 89

RESULT 21
US-10-369-493-5252

Sequence 5252, Application US/10369493
Publication No. US20030233673A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianning
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 5252
LENGTH: 1355
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-10-369-493-5252

Query Match
Best Local Similarity 100.0%; Score 11; DB 4; Length 1355;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 VOKKTFTKWN 52
|||||
Db 79 VOKKTFTKWN 89

RESULT 22
US-10-408-765A-741

Sequence 741, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Boia D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary W.
APPLICANT: Warnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
FILE REFERENCE: 660088,465
CURRENT APPLICATION NUMBER: US/10/408,765A
CURRENT FILING DATE: 2003-04-04
NUMBER OF SEQ ID NOS: 3077
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 741
LENGTH: 5373
TYPE: PRT
ORGANISM: Homo sapiens
US-10-408-765A-741

Query Match
Best Local Similarity 100.0%; Score 11; DB 4; Length 5373;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 42 VOKKTFTKWN 52
|||||
Db 44 VOKKTFTKWN 54

RESULT 23
US-10-805-684-151

Sequence 151, Application US/10805684
Publication No. US20050100966A1
GENERAL INFORMATION:
APPLICANT: SAKAMOTO, TAKESHI
APPLICANT: TAKEEDA, SHIZU
TITLE OF INVENTION: PHOS--INTERACTING PROTEINS AND USE THEREOF
FILE REFERENCE: 58748(70342)
CURRENT APPLICATION NUMBER: US/10/805,684
CURRENT FILING DATE: 2004-03-19
PRIOR APPLICATION NUMBER: 60/455,766
PRIOR FILING DATE: 2003-03-19
PRIOR APPLICATION NUMBER: 60/459,936
PRIOR FILING DATE: 2003-04-02
PRIOR APPLICATION NUMBER: 60/460,103
PRIOR FILING DATE: 2003-04-02
NUMBER OF SEQ ID NOS: 161
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 151
LENGTH: 5430
TYPE: PRT
ORGANISM: Homo sapiens
US-10-805-684-151

Query Match
Best Local Similarity 100.0%; Score 11; DB 5; Length 5430;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 VOKKTFTKWN 52
|||||
Db 79 VOKKTFTKWN 89

RESULT 24
US-11-097-143-40200

Sequence 40200, Application US/11097143
Publication No. US20050208558A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 40200

LENGTH: 7201
TYPE: PRT
ORGANISM: DROSOPHILA
US-11-097-143-40200

Query Match 5.0%; Score 11; DB 6; Length 7201;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 LLEVLSEGLRP 88
DB 29 LLEVLSEGLRP 39

RESULT 25
US-11-097-143-26253
Sequence 26253, Application US/11097143
Publication No. US20050208558A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 26253
LENGTH: 5385
TYPE: PRT
ORGANISM: DROSOPHILA
US-11-097-143-26253

Query Match 4.5%; Score 10; DB 6; Length 5385;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 OKKFTKWN 52
DB 39 OKKFTKWN 48

RESULT 26
US-11-097-143-28275
Sequence 28275, Application US/11097143
Publication No. US20050208558A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
FILE REFERENCE: CL000728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04

PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28275
LENGTH: 5496
TYPE: PRT
ORGANISM: DROSOPHILA
US-11-097-143-28275

Query Match 4.5%; Score 10; DB 6; Length 5496;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 OKKFTKWN 52
DB 150 OKKFTKWN 159

RESULT 27
US-10-363-616-413
Sequence 413, Application US/10363616
Publication No. US2004004181A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-113 (793)
CURRENT APPLICATION NUMBER: US/10/363,616
CURRENT FILING DATE: 2003-03-03
PRIOR APPLICATION NUMBER: 09/654,935
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 490
SEQ ID NO 413
LENGTH: 67
TYPE: PRT
ORGANISM: Homo sapiens
US-10-363-616-413

Query Match 4.1%; Score 9; DB 4; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTKWN 50
DB 37 VOKKFTKWN 45

RESULT 28
US-10-450-763-42286
Sequence 42286, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30

PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/549,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 42286
LENGTH: 117
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (63)..(86)
OTHER INFORMATION: Actin-type actin-binding domain proteins domain identified
OTHER INFORMATION: by eMATRIX, accession number BL00019B, p-value=1.000e-16, raw scc
OTHER INFORMATION: of 13.34
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (35)..(115)
OTHER INFORMATION: Calponin homology (CH) domain identified by Pfam, accession
OTHER INFORMATION: name CH, E-value=0.07, Pfam score of 4.1
US-10-450-763-42286

Query Match 4.1%; Score 9; DB 5; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VQKTFPTKW 50
|||||||
DB 36 VQKTFPTKW 44

RESULT 29
US-10-425-115-352183
Sequence 352183, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 352183
LENGTH: 123
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_84355C.1.pep
US-10-425-115-352183

Query Match 4.1%; Score 9; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 QOEERERR 151
|||||||
DB 75 QOEERERR 83

RESULT 30
US-10-424-599-237993
Sequence 237993, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 237993
LENGTH: 202
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_56934C.1.pep
US-10-424-599-237993

Query Match 4.1%; Score 9; DB 4; Length 202;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 QOEERERR 151
|||||||
DB 156 QOEERERR 164

RESULT 31
US-10-425-114-60052
Sequence 60052, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53113)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 60052
LENGTH: 202
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3637-254-C10_Flt.pep
US-10-425-114-60052

Query Match 4.1%; Score 9; DB 4; Length 202;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 QOEERERR 151
|||||||
DB 154 QOEERERR 162

RESULT 32
US-10-425-114-65298
Sequence 65298, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53113)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28

```
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 65298
/ LENGTH: 205
/ TYPE: PRT
/ ORGANISM: Zea mays subsp. mexicana
/ FEATURE:
/ OTHER INFORMATION: Clone ID: UC-ZMROSCINT021B06_FLI.pep
US-10-425-114-65298

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 205;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 QOEERKR 151
DB 157 QOEERKR 165

RESULT 33
US-10-425-114-36959
/ Sequence 36959, Application US/10425114
/ Publication No. US20040034888A1
/ GENERAL INFORMATION:
/ APPLICANT: Liu, Jingdong
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Screen, Steven E
/ APPLICANT: Tabaka, Jack E
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 36959
/ LENGTH: 240
/ TYPE: PRT
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB1170-019-F4_FLI.pep
US-10-425-114-36959

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 240;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 QOEERKR 151
DB 194 QOEERKR 202

RESULT 34
US-10-408-765A-731
/ Sequence 731, Application US/10408765A
/ Publication No. US20040101874A1
/ GENERAL INFORMATION:
/ APPLICANT: Ghosh, Soumitra S.
/ APPLICANT: Fahy, Bojin D.
/ APPLICANT: Zhang, Bing
/ APPLICANT: Gibson, Bradford W.
/ APPLICANT: Taylor, Steven W.
/ APPLICANT: Glenn, Gary M.
/ APPLICANT: Warnock, Dale E.
/ TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
/ TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
/ FILE REFERENCE: 660088.465
/ CURRENT APPLICATION NUMBER: US/10/408,765A
/ CURRENT FILING DATE: 2003-04-04
/ NUMBER OF SEQ ID NOS: 3077
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 731
/ LENGTH: 3433
```

```
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-408-765A-731

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 3433;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VQKTFK 50
DB 32 VQKTFK 40

RESULT 35
US-10-408-765A-732
/ Sequence 732, Application US/10408765A
/ Publication No. US20040101874A1
/ GENERAL INFORMATION:
/ APPLICANT: Ghosh, Soumitra S.
/ APPLICANT: Fahy, Bojin D.
/ APPLICANT: Zhang, Bing
/ APPLICANT: Gibson, Bradford W.
/ APPLICANT: Taylor, Steven W.
/ APPLICANT: Glenn, Gary M.
/ APPLICANT: Warnock, Dale E.
/ TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
/ TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
/ FILE REFERENCE: 660088.465
/ CURRENT APPLICATION NUMBER: US/10/408,765A
/ CURRENT FILING DATE: 2003-04-04
/ NUMBER OF SEQ ID NOS: 3077
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 732
/ LENGTH: 3433
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-408-765A-732

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 3433;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VQKTFK 50
DB 32 VQKTFK 40

RESULT 36
US-10-756-149-5681
/ Sequence 5681, Application US/10756149
/ Publication No. US20050181375A1
/ GENERAL INFORMATION:
/ APPLICANT: Aziz, Natsaba
/ APPLICANT: Zlotnik, Albert
/ TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
/ TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
/ FILE REFERENCE: file
/ CURRENT APPLICATION NUMBER: US/10/756,149
/ CURRENT FILING DATE: 2004-01-12
/ NUMBER OF SEQ ID NOS: 5818
/ SOFTWARE: Patencin version 3.2
/ SEQ ID NO 5681
/ LENGTH: 3433
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-10-756-149-5681

Query Match
Best Local Similarity 100.0%; Score 9; DB 5; Length 3433;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VQKTFK 50
```


Db 32 VQKFTTKW 40

RESULT 37

US-10-767-701-50379

Sequence 50379, Application US/10767701

Publication No. US20040172684A1

GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

FILE REFERENCE: 38-21(53535)B

CURRENT APPLICATION NUMBER: US/10/767,701

CURRENT FILING DATE: 2004-01-29

NUMBER OF SEQ ID NOS: 63128

SEQ ID NO 50379

LENGTH: 65

TYPE: PRT

ORGANISM: Sorghum bicolor

FEATURE:

OTHER INFORMATION: Clone ID: LIB3480-003-Pl-K1-Pl.pap

US-10-767-701-50379

Query Match 3.6%; Score 8; DB 4; Length 65;

Best Local Similarity 100.0%; Pred. No. 15;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 144 QEEERKR 151

Db 18 QEEERKR 25

RESULT 38

US-10-425-115-368921

Sequence 368921, Application US/10425115

Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 368921

LENGTH: 177

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(177)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577_99631C.1.pap

US-10-425-115-368921

Query Match 3.6%; Score 8; DB 4; Length 177;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 SGERLPKP 90

Db 93 SGERLPKP 100

RESULT 39

US-10-425-115-196724

Sequence 196724, Application US/10425115

Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 196724

LENGTH: 209

TYPE: PRT

ORGANISM: Zea mays

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(209)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577_110996C.1.pap

US-10-425-115-196724

Query Match 3.6%; Score 8; DB 4; Length 209;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 TKVSEEAR 167

Db 130 TKVSEEAR 137

RESULT 40

US-10-437-963-196932

Sequence 196932, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 196932

LENGTH: 213

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_92739C.1.pap

US-10-437-963-196932

Query Match 3.6%; Score 8; DB 4; Length 213;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 RQEEER 149

Db 126 RQEEER 133

RESULT 41

US-10-732-923-18430

Sequence 18430, Application US/10732923

Publication No. US20050108791A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Edgerton, Michael D
/ TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
/ FILE REFERENCE: 38-15(52796)C
/ CURRENT APPLICATION NUMBER: US/10/732,923
/ CURRENT FILING DATE: 2003-12-10
/ PRIOR APPLICATION NUMBER: 10/310,154
/ PRIOR FILING DATE: 2002-12-04
/ NUMBER OF SEQ ID NOS: 24149
/ SEQ ID NO 18430
/ LENGTH: 234
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-10-732-923-18430

Query Match
Best Local Similarity 100.0%; DB 5; Length 234;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 144 OEERERKR 151
Db 190 OEERERKR 197

RESULT 42
US-10-437-963-138843
/ Sequence 138843, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 138843
/ LENGTH: 246
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_40194C.1.pep
US-10-437-963-138843

Query Match
Best Local Similarity 100.0%; DB 4; Length 246;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 RROQEEER 148
Db 66 RROQEEER 73

RESULT 43
US-10-128-714-3119
/ Sequence 3119, Application US/10128714
/ Publication No. US20030119013A1
/ GENERAL INFORMATION:
/ APPLICANT: Jiang, Bo
/ APPLICANT: Hu, Wengqi
/ APPLICANT: Tishkoff, Daniel
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Eroszhkin, Alexey M
/ APPLICANT: Lemieux, Sebastien M
/ TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
/ TITLE OF INVENTION: Methods of Use
```

```
/ FILE REFERENCE: 10182-018-999
/ CURRENT APPLICATION NUMBER: US/10/128,714
/ CURRENT FILING DATE: 2002-04-23
/ PRIOR APPLICATION NUMBER: US 60/285,697
/ PRIOR FILING DATE: 2001-04-23
/ PRIOR APPLICATION NUMBER: US 60/287,066
/ PRIOR FILING DATE: 2001-04-27
/ PRIOR APPLICATION NUMBER: US 60/295,890
/ PRIOR FILING DATE: 2001-06-05
/ PRIOR APPLICATION NUMBER: US 60/303,899
/ PRIOR FILING DATE: 2001-07-09
/ PRIOR APPLICATION NUMBER: US 60/316,362
/ PRIOR FILING DATE: 2001-08-31
/ NUMBER OF SEQ ID NOS: 8603
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 3119
/ LENGTH: 358
/ TYPE: PRT
/ ORGANISM: Aspergillus fumigatus
US-10-128-714-3119

Query Match
Best Local Similarity 100.0%; DB 4; Length 358;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 146 BEERRRRP 153
Db 258 BEERRRRP 265

RESULT 44
US-10-437-963-140721
/ Sequence 140721, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 140721
/ LENGTH: 369
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_41894C.1.pep
US-10-437-963-140721

Query Match
Best Local Similarity 100.0%; DB 4; Length 369;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 141 RROQEEER 148
Db 156 RROQEEER 163

RESULT 45
US-10-425-115-242617
/ Sequence 242617, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
```

APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 242617
LENGTH: 382
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(382)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_152851C.1.pep
US-10-425-115-242617

Query Match
Best Local Similarity 3.6%; Score 8; DB 4; Length 382;
Matches 8; Conservativity 100.0%; Pred. No. 68;
Mismatches 0; Indels 0; Gaps 0;

Qy 160 TKVSEAE 167
|||||
Db 303 TKVSEAE 310

RESULT 46
US-10-128-714-8119
Sequence 8119, Application US/10128714
Publication No. US20030119013A1
GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wengqi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroshkin, Alexey M
APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
TITLE OF INVENTION: Methods of Use
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8119
LENGTH: 517
TYPE: PRT
ORGANISM: Aspergillus fumigatus
US-10-128-714-8119

Query Match
Best Local Similarity 3.6%; Score 8; DB 4; Length 517;
Matches 8; Conservativity 100.0%; Pred. No. 88;
Mismatches 0; Indels 0; Gaps 0;

Qy 146 EBERKRP 153
|||||
Db 417 EBERKRP 424

RESULT 47

US-11-097-143-16161
Sequence 16161, Application US/11097143
Publication No. US20050208558A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
TITLE OF INVENTION: DROSOPHILA GENES.
FILE REFERENCE: CLO00728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637
PRIOR FILING DATE: 2000-03-23
NUMBER OF SEQ ID NOS: 43008
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16161
LENGTH: 547
TYPE: PRT
ORGANISM: DROSOPHILA
US-11-097-143-16161

Query Match
Best Local Similarity 3.6%; Score 8; DB 6; Length 547;
Matches 8; Conservativity 100.0%; Pred. No. 93;
Mismatches 0; Indels 0; Gaps 0;

Qy 43 OKTFTKW 50
|||||
Db 21 OKTFTKW 28

RESULT 48
US-11-097-143-16218
Sequence 16218, Application US/11097143
Publication No. US20050208558A1
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig
APPLICANT: et al.
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
TITLE OF INVENTION: DROSOPHILA GENES.
FILE REFERENCE: CLO00728
CURRENT APPLICATION NUMBER: US/11/097,143
CURRENT FILING DATE: 2005-04-04
PRIOR APPLICATION NUMBER: 60/157,832
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: 60/160,191
PRIOR FILING DATE: 1999-10-19
PRIOR APPLICATION NUMBER: 60/161,932
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/164,769
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/173,383
PRIOR FILING DATE: 1999-12-28
PRIOR APPLICATION NUMBER: 60/175,693
PRIOR FILING DATE: 2000-01-12
PRIOR APPLICATION NUMBER: 60/184,831
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: 60/191,637

;; PRIOR FILING DATE: 2000-03-23
;; NUMBER OF SEQ ID NOS: 43008
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 16218
;; LENGTH: 690
;; TYPE: PRT
;; ORGANISM: DROSOPHILA
US-11-097-143-16218

Query Match 3.6%; Score 8; DB 6; Length 690;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 EEEERKR 152
|||
DB 180 EEEERKR 187

RESULT 49
US-10-733-969A-5
; Sequence 5, Application US/10733969A
; Publication No. US20040219572A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, JIE
; APPLICANT: HU, LIPING
; APPLICANT: LIU, TONG HUA
; APPLICANT: LU, ZHAO HUI
; APPLICANT: SHEN, YAN
; TITLE OF INVENTION: SPECIFIC MARKERS FOR PANCREATIC CANCER
; FILE REFERENCE: 21525
; CURRENT APPLICATION NUMBER: US/10/733,969A
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: EP 02026058.2
; PRIOR FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 911
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Alpha-actinin 4; Accession NO: 043707
US-10-733-969A-5

Query Match 3.6%; Score 8; DB 5; Length 911;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 SGERLKP 90
|||
DB 92 SGERLKP 99

RESULT 50
US-10-369-493-18546
; Sequence 18546, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18546
; LENGTH: 1190

;; TYPE: PRT
;; ORGANISM: Halobacterium sp. NRC-1
US-10-369-493-18546

Query Match 3.6%; Score 8; DB 4; Length 1190;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 LQDEREAV 42
|||
DB 768 LQDEREAV 775

Search completed: March 4, 2006, 07:14:28
Job time : 181 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2006, 07:06:31 ; Search time 40 Seconds
(without alignments)
529.192 Million cell updates/sec

Title: US-10-695-994A-7

Perfect score: 220

Sequence: 1 MELQRTSSISGPISPATYGTQ.....QNYKNFNSRRTASHDSWGM 220

Scoring table: ~~Clustal~~
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word-size: 6

Total number of hits satisfying chosen parameters: 1100

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 1000 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	44.1	2364	1 A44159	spectrin beta-G ch
2	38	17.3	2128	2 I52577	beta-spectrin - mo
3	38	17.3	2137	1 SUHUB	spectrin beta chai
4	27	12.3	2388	2 JEO271	beta spectrin, bet
5	19	8.6	2291	1 A46147	spectrin beta chai
6	17	7.7	30	2 S21195	spectrin beta chai
7	16	7.3	707	2 S29854	spectrin beta chai
8	14	6.4	2326	2 T29140	spectrin beta chai
9	11	5.0	41	2 I65645	hypothetical prote
10	11	5.0	101	2 I65644	Duchenne muscular
11	11	5.0	214	2 F59404	hypothetical prote
12	11	5.0	385	2 S66292	actin isoform pl
13	11	5.0	964	2 D58404	actin-croslinking
14	11	5.0	1355	2 T22552	plectin isoform pl
15	11	5.0	1825	2 T42725	hypothetical prote
16	11	5.0	1885	2 T30847	actin binding prot
17	11	5.0	3685	1 A27605	actin binding prot
18	11	5.0	4574	2 G02520	dyetrophin, musc
19	10	4.5	2396	2 T13714	plectin - human
20	9	4.1	98	2 S06461	Kakapo gene protei
21	9	4.1	230	2 J39161	dyetrophin, brain
22	9	4.1	292	2 JEO233	dyetrophin isoform 2
23	9	4.1	461	2 I39160	dyetrophin-1 - scal
24	9	4.1	511	2 I49290	dyetrophin isoform 1
25	9	4.1	602	2 C75120	dyetrophin - mouse
26	9	4.1	686	2 I49298	hypothetical prote
27	9	4.1	874	2 T15570	dyetrophin isoform 2
28	9	4.1	3433	1 S28381	hypothetical prote
29	9	4.1	3660	1 S02041	utrophin - human
					dyetrophin, musc

30	9	4.1	3678	2 S28916	dyetrophin - mouse
31	8	3.6	592	2 P66242	unknown protein, 9
32	8	3.6	894	1 PAMU2	alpha-actinin 2 -
33	8	3.6	897	2 S02032	alpha-actinin 2, 8
34	8	3.6	904	2 S45673	alpha-actinin, 115
35	8	3.6	911	2 UC7186	alpha-actinin-4 -
36	8	3.6	1089	2 T36663	protein kinase, tr
37	8	3.6	1190	2 E84193	chromosome segrega
38	8	3.6	1645	2 A37792	spectrin beta-H ch
39	8	3.6	1846	2 T42047	insulin receptor h
40	8	3.6	2414	2 A54277	transcription adap
41	8	3.6	4063	2 T42993	probable spectrin
42	8	3.6	4101	2 T23630	hypothetical prote
43	7	3.2	125	2 I65646	Duchenne muscular
44	7	3.2	172	2 S56905	hypothetical prote
45	7	3.2	218	2 B84165	glucose-1-phosphat
46	7	3.2	226	2 T49719	hypothetical prote
47	7	3.2	251	2 AB0957	conserved hypotet
48	7	3.2	266	1 P69844	conserved hypotet
49	7	3.2	279	2 G97246	arac-type DNA-bind
50	7	3.2	290	2 G86155	hypothetical prote
51	7	3.2	291	2 A85017	hypothetical prote
52	7	3.2	292	2 AG2127	hypothetical prote
53	7	3.2	311	2 T35159	hypothetical prote
54	7	3.2	314	2 JC4951	tropomycin T - scal
55	7	3.2	327	2 AH2368	hypothetical prote
56	7	3.2	356	2 S32147	natingenin 3-dioxy
57	7	3.2	357	2 S38338	natingenin 3-dioxy
58	7	3.2	357	2 S71772	natingenin 3-dioxy
59	7	3.2	357	2 S61415	natingenin 3-dioxy
60	7	3.2	357	2 A72220	conserved hypotet
61	7	3.2	358	2 T45755	flavonone 3-hydrox
62	7	3.2	360	2 S50830	flavonone 3-hydrox
63	7	3.2	397	2 A11201	Machado-Joseph dis
64	7	3.2	397	2 AD1559	glycine betaine AB
65	7	3.2	408	2 S63528	phosphoglycerate k
66	7	3.2	412	2 T02725	phosphoglycerate k
67	7	3.2	413	1 S28602	probable serine/ch
68	7	3.2	415	2 H80198	translation releas
69	7	3.2	433	2 T24351	phosphoglycerate k
70	7	3.2	470	2 D85111	hypothetical prote
71	7	3.2	478	2 B90507	hypothetical prote
72	7	3.2	481	2 S47441	gamma-glutamyltran
73	7	3.2	498	1 S67859	transport protein
74	7	3.2	515	2 T24781	Gund protein - xan
75	7	3.2	520	2 T22971	hypothetical prote
76	7	3.2	538	2 S67766	hypothetical prote
77	7	3.2	548	2 C75466	RNA-export mediato
78	7	3.2	569	2 T02435	phytoene dehydroge
79	7	3.2	595	2 C82960	probable oxysterol
80	7	3.2	621	1 T01935	probable two-compo
81	7	3.2	633	2 E95937	natingenin 3-dioxy
82	7	3.2	656	2 I49299	probable mureindep
83	7	3.2	669	2 C82984	receptor interacti
84	7	3.2	688	2 D72418	ATP-dependent DNA
85	7	3.2	701	2 H82978	DNA ligase - Therm
86	7	3.2	705	2 T01911	guanosine-3',5'-bi
87	7	3.2	714	2 A70983	hypothetical prote
88	7	3.2	721	2 A39707	probable fusa2 pro
89	7	3.2	755	2 H72262	erythrocyte membra
90	7	3.2	764	1 S14113	sensor histidine k
91	7	3.2	774	2 A24057	1-phosphatidylinos
92	7	3.2	819	2 A13197	glycophorin-bindin
93	7	3.2	822	2 S19947	Tomb-dependent rec
94	7	3.2	822	2 B49151	fibroblast growth
95	7	3.2	828	2 T52046	fibroblast growth
96	7	3.2	831	2 G82815	potassium channel
97	7	3.2	839	2 T50590	penicillin binding
98	7	3.2	848	2 S44732	class I INCENP pro
99	7	3.2	854	2 T17288	bos23.5 protein -
100	7	3.2	857	1 S05943	hypothetical prote
101	7	3.2	876	2 T19246	gelation factor -
102	7	3.2	877	2 I50591	hypothetical prote
					class II INCENP pr

103	7	3.2	972	1	GNXSIV	genome polyprotein
104	7	3.2	972	2	T09624	genome polyprotein
105	7	3.2	1028	2	T01714	hypothetical prote
106	7	3.2	1069	2	T00377	KIA0642 prote in -
107	7	3.2	1074	2	T01906	hypothetical prote
108	7	3.2	1091	2	C95133	exonuclease RexB (
109	7	3.2	1091	2	G98001	second chain of ma
110	7	3.2	1106	2	S38783	integrin alpha cha
111	7	3.2	1135	2	I61186	alpha-7 integrin -
112	7	3.2	1142	2	T00022	BI20 prote in - hum
113	7	3.2	1166	2	H71609	hypothetical prote
114	7	3.2	1254	2	A54818	myosin VI [similar
115	7	3.2	1257	2	A88536	protein B0523.5 (I
116	7	3.2	1263	2	T15496	hypothetical prote
117	7	3.2	1265	2	A59299	unconventional myo
118	7	3.2	1336	2	S41794	SEC3 prote in - yea
119	7	3.2	1407	1	S28589	trichothyalin - rab
120	7	3.2	1898	1	A45973	trichothyalin - hum
121	7	3.2	2440	2	S39162	transcription coac
122	7	3.2	2441	2	S39161	CRBB-binding prote
123	7	3.2	2962	2	T19756	hypothetical prote
124	7	3.2	4684	2	A59404	plectin [imported]
125	7	3.2	4687	1	A33638	plectin - rat
126	6	2.7	20	2	S33001	hypothetical prote
127	6	2.7	30	2	A44010	kit-ligand (alterm
128	6	2.7	34	4	S78720	TyA prote in YER138
129	6	2.7	34	4	S78704	TyA prote in YBL107
130	6	2.7	37	2	A23617	conglutin delta-2
131	6	2.7	38	2	T70014	kallikrein - mouse
132	6	2.7	42	2	D47753	beta-defensin-13 -
133	6	2.7	44	2	T16453	hypothetical prote
134	6	2.7	48	2	T14966	hypothetical prote
135	6	2.7	52	2	AD1587	hypothetical prote
136	6	2.7	58	2	F69163	hypothetical prote
137	6	2.7	65	2	F89906	hypothetical prote
138	6	2.7	67	2	H90323	hypothetical prote
139	6	2.7	71	2	AF2766	hypothetical prote
140	6	2.7	80	2	S02067	glycerol kinase (B
141	6	2.7	81	2	E72650	hypothetical prote
142	6	2.7	91	2	T71950	serum amyloid A -
143	6	2.7	91	2	T37327	insulin homolog ce
144	6	2.7	93	2	B83263	hypothetical prote
145	6	2.7	94	2	AC1240	B. subtilis YlXr p
146	6	2.7	94	2	AG1602	B. subtilis YlXr p
147	6	2.7	95	2	T09383	vpr prote in - huma
148	6	2.7	96	2	T01670	vpr prote in - huma
149	6	2.7	96	2	T09444	vpr prote in - huma
150	6	2.7	96	2	S54380	vpr prote in - huma
151	6	2.7	97	1	D44001	intermediate filam
152	6	2.7	98	2	S44095	probable membrane
153	6	2.7	105	2	AE0408	cysteine prote in s
154	6	2.7	106	2	S13027	oryacystatin II -
155	6	2.7	107	2	A38375	exclunuclease ABC C
156	6	2.7	110	2	A42573	conserved hypotet
157	6	2.7	111	2	B69117	conserved hypotet
158	6	2.7	113	2	C89828	hypothetical prote
159	6	2.7	117	2	A84147	hypothetical prote
160	6	2.7	118	2	B90178	hypothetical prote
161	6	2.7	119	2	F42523	A-ORF-A prote in -
162	6	2.7	122	2	A23521	serum amyloid A3 p
163	6	2.7	122	2	A32807	replication term in
164	6	2.7	122	2	AG0221	conserved hypotet
165	6	2.7	123	2	T16234	hypothetical prote
166	6	2.7	125	2	T05609	hypothetical prote
167	6	2.7	127	2	PC2268	hypothetical prote
168	6	2.7	127	2	B83157	cytochrome P450 pr
169	6	2.7	131	2	B96566	hypothetical prote
170	6	2.7	132	2	S77840	hypothetical prote
171	6	2.7	134	2	B64062	probable leucine-t
172	6	2.7	134	2	E57233	opacity-associated
173	6	2.7	134	2	C57233	complex in II - hum
174	6	2.7	134	2	JC4226	complex in II - rat
175	6	2.7	134	2	D57233	synaphin - bovine
						complex in II - mou
176	6	2.7	136	2	G72670	probable ribosomal
177	6	2.7	139	2	G75127	hypothetical prote
178	6	2.7	139	2	E71003	hypothetical prote
179	6	2.7	141	2	G82188	hypothetical prote
180	6	2.7	142	2	T18098	dCMP deaminase hom
181	6	2.7	144	2	C71149	hypothetical prote
182	6	2.7	144	2	T18867	hypothetical prote
183	6	2.7	145	2	A87574	hypothetical prote
184	6	2.7	146	2	A97166	flagellar prote in
185	6	2.7	145	2	S11398	pts fructose-speci
186	6	2.7	150	2	G64343	hypothetical prote
187	6	2.7	150	2	H87668	hypothetical prote
188	6	2.7	153	2	A33090	conglutin delta pr
189	6	2.7	159	2	H72613	hypothetical prote
190	6	2.7	161	2	D87406	ribosomal prote in
191	6	2.7	164	2	T16168	hypothetical prote
192	6	2.7	165	2	F69819	conserved hypotet
193	6	2.7	167	1	F45390	trans-regulatory s
194	6	2.7	167	1	VKLJVA	trans-regulatory s
195	6	2.7	167	1	VKLJVS	trans-regulatory s
196	6	2.7	167	2	S65670	phosphoprote in pho
197	6	2.7	170	2	G87153	50S ribosomal prot
198	6	2.7	171	2	G58930	hypothetical prote
199	6	2.7	172	2	A75592	hypothetical prote
200	6	2.7	172	2	F85745	unknown prote in en
201	6	2.7	173	2	T02332	probable Hmg-prote
202	6	2.7	176	2	S43463	mannose-binding le
203	6	2.7	176	2	F71803	hypothetical prote
204	6	2.7	176	2	D71336	probable peptidyl-
205	6	2.7	180	2	E70565	probable ribosomal
206	6	2.7	182	2	A83730	hypothetical prote
207	6	2.7	184	2	AG0736	hypothetical prote
208	6	2.7	184	2	F90873	probable antilelmi
209	6	2.7	185	2	C71899	hypothetical prote
210	6	2.7	185	2	AB3267	probable transcrip
211	6	2.7	190	2	AP2265	hypothetical prote
212	6	2.7	191	2	G69062	phenylacrylic acid
213	6	2.7	191	2	A55228	lme prote in homolo
214	6	2.7	192	2	S17812	hypothetical prote
215	6	2.7	193	2	F69756	cellulurium reaslan
216	6	2.7	193	2	S32992	hypothetical prote
217	6	2.7	193	2	H84531	hypothetical prote
218	6	2.7	194	2	AI2364	hypothetical prote
219	6	2.7	195	1	A64050	dCMP deaminase (BC
220	6	2.7	195	1	SAV1DS	delta large antige
221	6	2.7	195	1	C97509	hypothetical prote
222	6	2.7	195	2	C69839	3'-phosphadenosin
223	6	2.7	196	2	B84272	type 1 fibrinolyse
224	6	2.7	197	2	R8ECFB	hypothetical prote
225	6	2.7	198	2	H91287	hypothetical prote
226	6	2.7	198	2	C86129	phosphoglucosate d
227	6	2.7	198	2	T15353	hypothetical prote
228	6	2.7	198	2	S07397	lipocalin - mouse
229	6	2.7	200	2	T50367	hypothetical prote
230	6	2.7	203	2	T44588	DNA-directed RNA p
231	6	2.7	204	2	PS0262	adenylate kinase (
232	6	2.7	206	2	G70307	acyl-carrier prote
233	6	2.7	208	2	AC1151	acyl-carrier prote
234	6	2.7	208	2	AD1510	hypothetical prote
235	6	2.7	212	2	T10230	glycinin chain A7
236	6	2.7	212	2	JA0152	probable transcrip
237	6	2.7	213	2	T44588	conserved hypotet
238	6	2.7	214	2	C64425	GTP cyclonhydrolase
239	6	2.7	216	2	AI2395	RNase III - Aquife
240	6	2.7	221	2	G70381	hypothetical prote
241	6	2.7	221	2	C96918	GTP-binding protel
242	6	2.7	223	2	S42679	probable cation tr
243	6	2.7	225	2	T36651	hypothetical prote
244	6	2.7	226	2	D83875	nitroreductase, bl
245	6	2.7	227	2	B90429	ribonuclease II (B
246	6	2.7	229	2	S66342	two-component reas
247	6	2.7	229	2	G70073	cytidylate kinase
248	6	2.7	229	2	C82141	

249	6	2.7	230	2	B83879	hypothetical prote
250	6	2.7	230	2	A75260	hypothetical prote
251	6	2.7	231	2	G81026	inositol monophosp
252	6	2.7	232	2	B63392	ABC transporter, A
253	6	2.7	232	2	T43754	hypothetical prote
254	6	2.7	233	1	WMBEHK	UL3 protein - huma
255	6	2.7	233	1	D69630	menaquinone biosyn
256	6	2.7	233	1	F69178	conserved hypothet
257	6	2.7	236	2	A10981	conserved hypothet
258	6	2.7	236	2	A75431	hypothetical prote
259	6	2.7	237	2	A99103	26S proteasome 10T
260	6	2.7	237	2	AE3619	transcription regu
261	6	2.7	238	1	QCBF74	US4 protein - huma
262	6	2.7	240	2	T03379	heat shock protein
263	6	2.7	240	2	G87345	glutathione S-tran
264	6	2.7	240	2	T41254	probable coactmer d
265	6	2.7	240	2	T24239	hypothetical prote
266	6	2.7	241	2	A87551	glutathione S-tran
267	6	2.7	241	2	S44893	ZK1236.7 protein -
268	6	2.7	241	2	S49014	lysin - lactococcu
269	6	2.7	241	2	C86492	hypothetical prote
270	6	2.7	241	2	S61925	cyst wall protein
271	6	2.7	241	2	C72130	hypothetical prote
272	6	2.7	242	2	T36929	hypothetical prote
273	6	2.7	243	2	T46221	PRE-MRNA SPLICING
274	6	2.7	244	2	S44822	F4AE2.3 protein -
275	6	2.7	244	2	T44624	cinI protein (limp
276	6	2.7	246	2	T16403	hypothetical prote
277	6	2.7	247	2	A87174	hypothetical prote
278	6	2.7	247	2	T37001	hypothetical prote
279	6	2.7	249	2	AC0698	probable pathogeni
280	6	2.7	250	2	C64304	ABC transporter su
281	6	2.7	251	2	H75586	probable urea/short
282	6	2.7	252	1	S53047	indole-3-glycerol
283	6	2.7	252	2	B83584	probable molybdenu
284	6	2.7	252	2	T12795	hypothetical prote
285	6	2.7	252	2	A45594	ORF 5' of calmodul
286	6	2.7	253	1	A40582	type IV prepillin p
287	6	2.7	253	1	D82273	leader peptidase I
288	6	2.7	255	1	S04899	my-related protei
289	6	2.7	255	2	T36778	probable enoyl-(ac
290	6	2.7	256	2	A49156	acetyl-CoA carboxy
291	6	2.7	257	2	B83050	probable transcrip
292	6	2.7	257	2	T12092	G-box-binding prot
293	6	2.7	259	2	T21715	hypothetical prote
294	6	2.7	260	2	S11562	probable MASH-1 pr
295	6	2.7	261	2	T24532	hypothetical prote
296	6	2.7	264	2	D81971	glycinin A5A4B3 pr
297	6	2.7	268	2	P00199	dihydrodipicolinat
298	6	2.7	269	2	H84946	amino acid ABC tra
299	6	2.7	269	2	AF1737	hypothetical prote
300	6	2.7	269	2	C84241	pyruvate dehydroge
301	6	2.7	270	2	A13642	hypothetical prote
302	6	2.7	270	2	T22213	hypothetical prote
303	6	2.7	271	2	S67072	probable membrane
304	6	2.7	272	2	A54617	transcription fact
305	6	2.7	272	2	AF0595	conserved hypothet
306	6	2.7	272	2	S73649	triacylglycerol 1i
307	6	2.7	276	2	C86030	hypothetical prote
308	6	2.7	276	2	G91183	hypothetical prote
309	6	2.7	276	2	S47775	ylaf protein - Bac
310	6	2.7	276	2	T33529	hypothetical prote
311	6	2.7	281	2	H97124	hypothetical prote
312	6	2.7	283	2	G84861	hypothetical prote
313	6	2.7	284	2	D95109	conserved hypothet
314	6	2.7	285	2	S21562	hypothetical prote
315	6	2.7	285	2	G66875	hypothetical prote
316	6	2.7	286	2	S17548	alpha-actinin - ra
317	6	2.7	286	2	G65082	putative general s
318	6	2.7	287	1	PWBY	inorganic diphosph
319	6	2.7	287	1	PWVKL	hypothetical prote
320	6	2.7	287	2	T22325	hypothetical prote
321	6	2.7	287	2	D69536	hypothetical prote
322	6	2.7	287	2	T05338	hypothetical prote
323	6	2.7	288	2	F64448	methylyliogen-red
324	6	2.7	288	2	S58219	ABA-inducible prot
325	6	2.7	288	2	H88109	protein T24E12.3 l
326	6	2.7	289	2	T03395	probable lipase -
327	6	2.7	290	2	S39854	trax protein - Str
328	6	2.7	290	2	JC2097	legumin type B alp
329	6	2.7	291	1	JH0489	phosphoribosylamin
330	6	2.7	291	2	S55291	phosphoribosylamin
331	6	2.7	291	2	S55292	phosphoribosylamin
332	6	2.7	291	2	S55293	phosphoribosylamin
333	6	2.7	291	2	H70678	hypothetical prote
334	6	2.7	291	2	AH1308	methylyltransferase
335	6	2.7	291	2	AH1680	methylyltransferase
336	6	2.7	291	2	S27721	hypothetical prote
337	6	2.7	291	2	T40277	probable membrane
338	6	2.7	294	2	S59774	conserved membrane
339	6	2.7	297	2	G97977	conserved hypothet
340	6	2.7	298	2	S32368	beta-SNAP protein
341	6	2.7	298	2	H87533	peptidase, M23/M37
342	6	2.7	300	2	P80274	hypothetical prote
343	6	2.7	300	2	AC1924	hypothetical prote
344	6	2.7	301	2	B85087	probable phosphogl
345	6	2.7	302	2	A33964	regulatory protein
346	6	2.7	303	2	JC5140	urate oxidase (BC
347	6	2.7	303	2	T29321	hypothetical prote
348	6	2.7	303	2	S77235	hypothetical prote
349	6	2.7	305	2	E70410	hypothetical prote
350	6	2.7	307	2	T29222	hypothetical prote
351	6	2.7	307	2	AG2862	conserved hypothet
352	6	2.7	307	2	P97639	hypothetical prote
353	6	2.7	308	2	D71353	hypothetical prote
354	6	2.7	308	2	T46294	hypothetical prote
355	6	2.7	308	2	A29379	glycine-rich prote
356	6	2.7	310	2	T05733	ribosome-associate
357	6	2.7	310	2	T34962	hypothetical prote
358	6	2.7	310	2	AB2235	hypothetical prote
359	6	2.7	310	2	T46429	hypothetical prote
360	6	2.7	311	2	AF2064	hypothetical prote
361	6	2.7	311	2	F84923	pale crease protein
362	6	2.7	315	2	AH0897	probable carboxydr
363	6	2.7	315	2	S76267	hypothetical prote
364	6	2.7	315	2	T45099	methylyltetrahydro
365	6	2.7	316	2	S58719	probable membrane
366	6	2.7	317	2	A11887	coproporphyrinogen
367	6	2.7	317	2	B37804	feer protein - Esc
368	6	2.7	317	2	F86901	hypothetical prote
369	6	2.7	317	2	JC2110	tropomyosin-relate
370	6	2.7	319	2	B87365	hypothetical prote
371	6	2.7	320	2	S78604	taurin-binding pro
372	6	2.7	320	2	C90681	taurine transport
373	6	2.7	320	2	G85531	taurine transport
374	6	2.7	320	2	S18324	lamin C - mouse (f
375	6	2.7	321	2	T47138	hypothetical prote
376	6	2.7	322	2	H69457	ornithine cyclodex
377	6	2.7	322	2	S75280	hypothetical prote
378	6	2.7	322	2	AF3047	conserved hypothet
379	6	2.7	323	2	E71168	hypothetical prote
380	6	2.7	324	2	T00939	hypothetical prote
381	6	2.7	325	2	D75466	hypothetical prote
382	6	2.7	325	2	S23007	pyruvate synthase
383	6	2.7	326	2	AD3211	CAMP response elem
384	6	2.7	327	1	RDVZAS	transcription regu
385	6	2.7	327	1	E98238	ribonucleoside-dip
386	6	2.7	328	2	S22299	CAMP response elem
387	6	2.7	328	2	A99416	alcohol dehydrogen
388	6	2.7	328	2	A35663	CAMP response elem
389	6	2.7	328	2	C90196	conserved hypothet
390	6	2.7	328	2	T48248	ABC transporter-1i
391	6	2.7	329	2	T42996	hypothetical prote
392	6	2.7	329	2	S07577	legumin storage pr
393	6	2.7	330	2	T01190	DNA-binding protei
394	6	2.7	332	2	H82340	gluconate utilizat

395	6	2.7	333	2	AD2608	aldo-keto reductas
396	6	2.7	333	2	UC513	UDP-glucose 4-epime
397	6	2.7	334	2	AB0643	DNA polymerase III
398	6	2.7	334	2	B81406	probable integral
399	6	2.7	335	2	S44224	scyllo-inosamine-4
400	6	2.7	335	2	S44227	amidino transferase
401	6	2.7	335	2	S07576	legumin storage pr
402	6	2.7	335	2	S07576	legumin storage pr
403	6	2.7	336	2	AB2973	probable transcript
404	6	2.7	337	2	S42416	probable membrane
405	6	2.7	338	2	B83861	hypothetical prote
406	6	2.7	340	2	C70030	hypothetical prote
407	6	2.7	340	2	A69420	hypothetical prote
408	6	2.7	341	2	S03343	CAMP response elem
409	6	2.7	343	2	H84000	spore photoproduct
410	6	2.7	345	2	AB1813	hypothetical prote
411	6	2.7	345	2	T37685	probable vanadate
412	6	2.7	346	2	D47211	GTP-binding protei
413	6	2.7	347	2	S50403	TF134 protein - ye
414	6	2.7	347	2	G95402	hypothetical prote
415	6	2.7	348	2	E69162	sulfate transport
416	6	2.7	350	2	B82713	lipase modulator X
417	6	2.7	352	2	S68687	glucose-1-phosphat
418	6	2.7	352	2	C84342	phosphoribosyl tra
419	6	2.7	353	2	T46521	probable GTP-1-gi
420	6	2.7	353	2	C72479	hypothetical prote
421	6	2.7	353	2	S69047	hypothetical prote
422	6	2.7	354	2	A86843	prephenate dehydro
423	6	2.7	354	2	S28765	hypothetical prote
424	6	2.7	354	2	G81939	hypothetical prote
425	6	2.7	355	2	I49339	macrophage inflamm
426	6	2.7	355	2	H96012	probable sugar upt
427	6	2.7	359	2	T48396	fructose-bisphosph
428	6	2.7	359	2	AH1145	conserved hypothet
429	6	2.7	359	2	T47770	hypothetical prote
430	6	2.7	360	2	T22576	hypothetical prote
431	6	2.7	360	2	T35374	probable regulator
432	6	2.7	360	2	AF3175	DNA-directed DNA p
433	6	2.7	360	2	JCS804	CDP-glucose 4,6-deh
434	6	2.7	362	1	S61924	cyst wall protein -
435	6	2.7	364	2	S77360	chid protein - Syn
436	6	2.7	364	2	S28771	polysialacturonase
437	6	2.7	364	2	AB2993	glycosyl transferas
438	6	2.7	365	2	H90113	nucleolar snRNP pr
439	6	2.7	366	2	T50453	probable glycosylt
440	6	2.7	366	2	G83460	hypothetical prote
441	6	2.7	367	2	H82307	conserved hypothet
442	6	2.7	368	2	D70481	glucosyl transfera
443	6	2.7	368	2	C29356	hydroxyproline-ric
444	6	2.7	369	2	C97390	aldo/keto reductas
445	6	2.7	370	2	H64480	hypothetical prote
446	6	2.7	370	2	JC7591	spinal cord-deriva
447	6	2.7	374	2	F87596	hypothetical prote
448	6	2.7	374	2	E95167	ABC transporter, A
449	6	2.7	375	2	H84666	probable MYB famiI
450	6	2.7	378	2	S38962	serpin - pig
451	6	2.7	378	2	G02313	CDC37 homolog - hu
452	6	2.7	379	2	T36831	probable two-compo
453	6	2.7	379	2	P96565	hypothetical prote
454	6	2.7	380	2	T48953	UDP-glucose-hexose-
455	6	2.7	380	2	T48953	hypothetical prote
456	6	2.7	380	2	T26255	hypothetical prote
457	6	2.7	381	2	S65212	hypothetical prote
458	6	2.7	381	2	C84650	hypothetical prote
459	6	2.7	381	2	S69065	probable membrane
460	6	2.7	382	2	T15204	hypothetical prote
461	6	2.7	383	2	G84597	probable XAP-5 pro
462	6	2.7	383	2	C96581	hypothetical prote
463	6	2.7	385	1	RGBSDS	degradative enzyme
464	6	2.7	385	2	B83725	hypothetical prote
465	6	2.7	386	2	T02640	hypothetical prote
466	6	2.7	386	2	S37691	ran GTPase activat
467	6	2.7	386	2	T39461	hypothetical prote
468	6	2.7	387	2	A82036	fatty oxidation co
469	6	2.7	387	2	H86445	probable G-Box bin
470	6	2.7	387	2	AD1960	two-component hydr
471	6	2.7	388	2	A42465	integrase - phage
472	6	2.7	388	2	T41089	hypothetical prote
473	6	2.7	389	2	T03612	chalcone synthase
474	6	2.7	389	2	T02970	hypothetical prote
475	6	2.7	389	2	C96830	unknown protein F1
476	6	2.7	389	2	T27085	hypothetical prote
477	6	2.7	390	1	W2MLRB	E2 protein - cotto
478	6	2.7	390	2	S65672	phosphoprotein pho
479	6	2.7	391	2	G89822	hypothetical prote
480	6	2.7	391	2	B64307	ammonium transport
481	6	2.7	393	2	A10319	probable 2-octapre
482	6	2.7	393	2	I49016	phosphoprotein pho
483	6	2.7	394	2	E71136	threonine synthase
484	6	2.7	394	2	JQ2041	polymerase-associ
485	6	2.7	394	2	B75439	conserved hypothet
486	6	2.7	395	2	B90546	phosphopentomutase
487	6	2.7	395	2	B75512	conserved hypothet
488	6	2.7	396	2	I75615	mammary tumor inte
489	6	2.7	397	2	S66700	probable membrane
490	6	2.7	398	2	H72257	hypothetical prote
491	6	2.7	398	2	A80975	probable racemase
492	6	2.7	399	1	HMXRM2	sigma 1 protein -
493	6	2.7	399	2	G96690	unknown protein P2
494	6	2.7	399	2	B86289	TFEN1.7 protein -
495	6	2.7	400	2	C89955	hypothetical prote
496	6	2.7	400	2	H70405	conserved hypothet
497	6	2.7	401	2	D72386	threonine ammonia-
498	6	2.7	402	2	A72312	conserved hypothet
499	6	2.7	403	1	C69213	protein-export mem
500	6	2.7	404	2	AB1832	hypothetical prote
501	6	2.7	406	2	H87392	hypothetical prote
502	6	2.7	406	2	T24492	tyrosine-CRN A19a
503	6	2.7	407	2	C64250	bZIP DNA-binding p
504	6	2.7	407	2	T14909	bZIP DNA-binding p
505	6	2.7	408	2	T13254	biphényl dioxygens
506	6	2.7	408	2	F85023	probable potassium
507	6	2.7	408	2	H98290	hypothetical prote
508	6	2.7	409	2	T11901	NADH2 dehydrogenas
509	6	2.7	411	2	C98142	hypothetical prote
510	6	2.7	413	2	H88481	protein let-756 (I
511	6	2.7	419	2	AD3246	P-450 monooxygenas
512	6	2.7	420	2	T14911	bZIP DNA-binding p
513	6	2.7	420	2	JC7229	vitamin D receptor
514	6	2.7	420	2	S43559	collided coil protei
515	6	2.7	421	2	S35301	rflB protein - Yer
516	6	2.7	421	2	AH3627	maltose-binding pe
517	6	2.7	422	2	S35480	hypothetical prote
518	6	2.7	423	2	G95880	probable trehalose
519	6	2.7	423	2	AF3466	trehalose/maltose
520	6	2.7	424	2	AB1034	UV protection prot
521	6	2.7	424	2	B38176	saab protein - Sal
522	6	2.7	425	2	B64403	translacion releas
523	6	2.7	425	2	UC7230	vitamin D receptor
524	6	2.7	426	2	S67101	probable nicotinat
525	6	2.7	429	2	T51540	hypothetical prote
526	6	2.7	436	2	T14816	hypothetical prote
527	6	2.7	436	2	T15331	hypothetical prote
528	6	2.7	439	2	T49334	basic transcriptio
529	6	2.7	439	2	T31031	hypothetical prote
530	6	2.7	442	2	H69181	hypothetical prote
531	6	2.7	443	2	I39538	alpha-amyase - Ae
532	6	2.7	443	2	T47963	hypothetical prote
533	6	2.7	445	2	I40682	phosphogluconate d
534	6	2.7	445	2	I40685	phosphogluconate d
535	6	2.7	445	2	I40684	phosphogluconate d
536	6	2.7	445	2	I40681	phosphogluconate d
537	6	2.7	445	2	T31581	chromogranin A pre
538	6	2.7	446	2	A33284	hypothetical prote
539	6	2.7	447	2	AP2295	protein kinase, pr
540	6	2.7	448	2	F95122	

541	6	2.7	449	2	C69086	614	6	2.7	517	2	T51328	transcription init
542	6	2.7	451	2	G82218	615	6	2.7	518	2	AB0843	glutamate-cysteine
543	6	2.7	452	2	S64211	616	6	2.7	519	2	B82932	spermidine/putresc
544	6	2.7	452	2	AT2966	617	6	2.7	520	1	G86662	2',3'-cyclic-nucle
545	6	2.7	452	2	B98316	618	6	2.7	521	1	T37252	probable matrix me
546	6	2.7	452	2	T22948	619	6	2.7	523	1	DEF66	glucose-6-phosphat
547	6	2.7	454	2	D98033	620	6	2.7	524	2	A47740	glucose-6-phosphat
548	6	2.7	455	2	AD1723	621	6	2.7	525	2	T31563	hypothetical prote
549	6	2.7	455	2	S33073	622	6	2.7	526	2	T07082	lycopene epsilon-c
550	6	2.7	457	2	S39079	623	6	2.7	526	2	UC7248	alpha-N-acetylglata
551	6	2.7	458	1	C70058	624	6	2.7	527	1	CSRT	catalase (EC 1.11.
552	6	2.7	459	2	B64171	625	6	2.7	527	2	A36695	catalase (EC 1.11.
553	6	2.7	462	1	A60746	626	6	2.7	527	2	P64210	hypothetical prote
554	6	2.7	462	2	I54414	627	6	2.7	530	1	S57907	carboxypeptidase D
555	6	2.7	462	2	C34829	628	6	2.7	530	2	G95279	probable ABC trans
556	6	2.7	464	2	G86566	629	6	2.7	531	2	S48087	t-complex-type mol
557	6	2.7	464	2	D72058	630	6	2.7	531	2	S43063	t-complex-type mol
558	6	2.7	465	2	T40697	631	6	2.7	531	2	B69346	signal-transducin
559	6	2.7	465	2	B70213	632	6	2.7	532	2	C87793	protein C27A12.3 [
560	6	2.7	468	2	I62463	633	6	2.7	534	2	T39903	serine-rich protei
561	6	2.7	468	2	T34679	634	6	2.7	535	2	S19729	xylin 1,4-beta-xy
562	6	2.7	468	2	G71231	635	6	2.7	537	2	T48599	hypothetical prote
563	6	2.7	469	2	S69640	636	6	2.7	538	2	T40151	hypothetical prote
564	6	2.7	470	2	T10193	637	6	2.7	538	2	B84759	hypothetical prote
565	6	2.7	470	2	T22785	638	6	2.7	539	2	T15943	t-complex protein
566	6	2.7	470	2	T29380	639	6	2.7	540	2	S21825	vicillin-like stora
567	6	2.7	471	2	T21102	640	6	2.7	541	2	S43061	t-complex-type mol
568	6	2.7	471	2	S66829	641	6	2.7	545	2	T05510	hypothetical prote
569	6	2.7	471	2	S41768	642	6	2.7	546	1	SYBYHM	histidine-tRNA lig
570	6	2.7	473	2	A54691	643	6	2.7	546	1	S63106	hypothetical prote
571	6	2.7	474	2	T16441	644	6	2.7	551	1	JC5225	dsRNA-activated pr
572	6	2.7	475	1	Z28BL7	645	6	2.7	552	1	G69652	endopeptidase Ia (
573	6	2.7	476	2	T03745	646	6	2.7	553	2	G83864	transcription regu
574	6	2.7	478	1	S73920	647	6	2.7	559	2	S52941	period protein (Cl
575	6	2.7	478	1	S60754	648	6	2.7	560	2	S52943	period protein (Cl
576	6	2.7	478	2	S69974	649	6	2.7	560	2	S11004	glycinn G4 precu
577	6	2.7	479	2	D86192	650	6	2.7	561	2	AF0629	cell invasion prot
578	6	2.7	481	2	B75167	651	6	2.7	562	1	FMSYG5	glycinn chain A5A
579	6	2.7	481	2	H70679	652	6	2.7	562	2	T33174	hypothetical prote
580	6	2.7	481	2	F71048	653	6	2.7	562	2	PC7084	GTP-binding protei
581	6	2.7	481	2	S04605	654	6	2.7	562	2	S20946	glycinn G4 precu
582	6	2.7	484	2	AG3184	655	6	2.7	562	2	T49788	related to mezozi
583	6	2.7	484	2	AF3529	656	6	2.7	563	2	S54802	glycinn A5A4B3 ch
584	6	2.7	484	2	S11003	657	6	2.7	566	2	S22477	vicillin precursor
585	6	2.7	484	2	A24942	658	6	2.7	566	2	A84485	hypothetical prote
586	6	2.7	486	1	TVHUEG	659	6	2.7	567	2	H84770	probable receptor-
587	6	2.7	486	2	S74319	660	6	2.7	567	2	C69014	oxaloacetate decar
588	6	2.7	488	2	T05313	661	6	2.7	569	2	F87501	hypothetical prote
589	6	2.7	488	2	C86183	662	6	2.7	569	2	S74053	probable acylamino
590	6	2.7	492	2	A83632	663	6	2.7	571	2	T31170	matutase-related p
591	6	2.7	492	2	B72389	664	6	2.7	572	1	VBHULC	lamin C - human
592	6	2.7	494	2	T02523	665	6	2.7	573	2	A53234	globulin-1S, GLB1S
593	6	2.7	495	2	AB0636	666	6	2.7	574	2	S04333	gene a17 protein -
594	6	2.7	496	2	E88987	667	6	2.7	574	2	S48860	transporter bindin
595	6	2.7	496	2	AD1211	668	6	2.7	575	2	G69834	single-stranded DN
596	6	2.7	496	2	AD1567	669	6	2.7	577	2	D65073	sBDNA exonuclease
597	6	2.7	498	2	C47021	670	6	2.7	577	2	D51099	sBDNA exonuclease
598	6	2.7	498	2	F83523	671	6	2.7	577	2	H85944	hypothetical prote
599	6	2.7	499	1	Z2BP22	672	6	2.7	577	2	AD1440	probable membrane
600	6	2.7	499	2	F72768	673	6	2.7	577	2	S64250	hypothetical prote
601	6	2.7	499	2	F69833	674	6	2.7	578	2	S23847	probable oligopept
602	6	2.7	501	2	T00213	675	6	2.7	578	2	G75632	probable oligopept
603	6	2.7	501	2	S22669	676	6	2.7	579	2	UT0494	alpha-glucosidase
604	6	2.7	501	2	I39360	677	6	2.7	579	2	JC7170	fibxin-like 71 K
605	6	2.7	506	2	F71513	678	6	2.7	580	2	T21493	hypothetical prote
606	6	2.7	506	2	C81704	679	6	2.7	581	2	T45889	erin - bovine
607	6	2.7	507	2	AH2717	680	6	2.7	582	2	B53234	vicillin-like stora
608	6	2.7	507	2	C97499	681	6	2.7	582	2	JC7285	GTP-binding protei
609	6	2.7	509	2	AB0683	682	6	2.7	584	2	T40500	2-isopropylmalate
610	6	2.7	512	1	G69170	683	6	2.7	584	2	S53902	regulatory protein
611	6	2.7	513	1	H89805	684	6	2.7	584	2	F75090	archaeosine trna-r
612	6	2.7	516	1	FMSYG3	685	6	2.7	586	1	A34400	erin [validated]
613	6	2.7	516	2	AD2279	686	6	2.7	590	2	B86440	probable protein k

687	6	2.7	590	2	P82686	phosphotransferase
688	6	2.7	591	2	T14364	probable transcript
689	6	2.7	591	2	SS1303	hypothetical prote
690	6	2.7	599	2	D97818	nitrogen regulatio
691	6	2.7	600	2	T41165	hypothetical prote
692	6	2.7	600	2	T02692	hypothetical prote
693	6	2.7	601	2	G71666	nitrogen regulatio
694	6	2.7	604	2	S66993	hypothetical prote
695	6	2.7	604	2	AC1083	toxin components h
696	6	2.7	605	2	B87365	sensor histidine k
697	6	2.7	608	2	S72177	dextranase (EC 3.2
698	6	2.7	611	2	S65472	peptidyl-dipeptida
699	6	2.7	612	2	B81246	glutamine-fructose
700	6	2.7	612	2	T40936	probable leukotrie
701	6	2.7	615	2	AK3069	cold-shock dead-bo
702	6	2.7	615	2	P98217	inducible ATP-inde
703	6	2.7	616	2	A40595	metilmalonyl-CoA
704	6	2.7	619	2	A23872	2-isopropylmalate
705	6	2.7	619	2	D63194	PEP112-like protei
706	6	2.7	620	2	JH0821	95k golgi antigen
707	6	2.7	621	2	T48187	hypothetical prote
708	6	2.7	621	2	T37708	hypothetical prote
709	6	2.7	623	2	T40685	phosphatidylcholin
710	6	2.7	623	2	B82536	ABC transporter At
711	6	2.7	625	2	S68633	hypothetical prote
712	6	2.7	625	2	T132739	hypothetical prote
713	6	2.7	629	2	A29666	keratin, 65k type
714	6	2.7	630	2	T47177	hypothetical prote
715	6	2.7	633	2	H96748	unknown protein T1
716	6	2.7	634	2	B86293	T2418.1 protein -
717	6	2.7	635	2	D64831	ABC-type transport
718	6	2.7	635	2	G85621	hypothetical prote
719	6	2.7	635	2	A10625	ABC transporter At
720	6	2.7	635	2	A90758	hypothetical prote
721	6	2.7	637	2	H96592	probable multispin
722	6	2.7	637	2	AB0173	ABC transporter At
723	6	2.7	639	2	G82194	ABC transporter, A
724	6	2.7	640	2	T08758	hypothetical prote
725	6	2.7	640	2	H83267	probable ATP-Bindl
726	6	2.7	642	1	S29320	fibrin - Yeast (S
727	6	2.7	645	2	A47081	triacylglycerol li
728	6	2.7	646	2	T47154	hypothetical prote
729	6	2.7	646	2	T38171	probable serine/th
730	6	2.7	647	2	B64170	ABC-type transport
731	6	2.7	648	2	B40727	S-M checkpoint con
732	6	2.7	649	2	AF2866	methyl-accepting c
733	6	2.7	649	2	D85135	hypothetical prote
734	6	2.7	651	2	C69374	conserved hypothet
735	6	2.7	652	2	T28924	hypothetical prote
736	6	2.7	653	2	A82647	phage-related inte
737	6	2.7	657	2	B75484	glycosyl hydrolase
738	6	2.7	657	2	S05517	lamin - chicken
739	6	2.7	658	2	A11385	exonuclease ABC (
740	6	2.7	658	2	D84869	probable receptor
741	6	2.7	658	2	S68418	protein phosphatas
742	6	2.7	659	2	F70175	rep helicase, sing
743	6	2.7	659	2	S11736	resistance protein
744	6	2.7	659	2	S11737	resistance protein
745	6	2.7	659	2	D84633	probable multispin
746	6	2.7	662	2	AG0479	probable alpha-any
747	6	2.7	662	2	T18233	probable transcript
748	6	2.7	664	1	VEH01A	lamin A - human
749	6	2.7	664	2	T24459	hypothetical prote
750	6	2.7	664	2	C84869	probable receptor
751	6	2.7	665	2	S27267	lamin A - rat
752	6	2.7	665	2	S02358	lamin A - African
753	6	2.7	665	2	S28182	lamin A - mouse
754	6	2.7	666	2	T03090	hypothetical prote
755	6	2.7	666	2	H96723	hypothetical prote
756	6	2.7	669	2	T28754	hypothetical prote
757	6	2.7	670	2	C97643	similar to mcpo ge
758	6	2.7	670	2	D96739	hypothetical prote
759	6	2.7	672	2	A41075	inositol-1,4,5-tri
760	6	2.7	677	1	A45264	system b(0,+) amln
761	6	2.7	678	1	G75524	translaction elonga
762	6	2.7	683	1	A41785	system b(0,+) amln
763	6	2.7	683	2	D97011	sensory transducti
764	6	2.7	684	2	A56154	Abi substrate ena
765	6	2.7	685	1	SXBPT4	NAD+-protein ADP-r
766	6	2.7	685	1	EBYS2	suppressor 2 prote
767	6	2.7	687	1	BB949	beta-galactosidase
768	6	2.7	687	2	D96553	hypothetical prote
769	6	2.7	687	2	T29148	hypothetical prote
770	6	2.7	689	2	T08988	cadmium-transporti
771	6	2.7	693	2	C83821	glycyl-tRNA synthe
772	6	2.7	696	2	S39837	SKTS protein - Yea
773	6	2.7	699	2	B84565	hypothetical prote
774	6	2.7	705	2	T49461	probable croi prot
775	6	2.7	708	2	A72264	polynucleotide pho
776	6	2.7	711	2	A85352	cadmium-transporti
777	6	2.7	712	2	C71419	hypothetical prote
778	6	2.7	716	2	S45262	NF-AT component -
779	6	2.7	716	2	B90738	probably ATP-depen
780	6	2.7	716	2	G85588	probable ATP-depen
781	6	2.7	716	2	G64816	probable ATP-depen
782	6	2.7	718	2	A83282	probable TonB-depe
783	6	2.7	718	2	JC5805	transcription fact
784	6	2.7	722	2	H86986	endo-1,4-beta gluc
785	6	2.7	723	2	C36963	cellulose synthase
786	6	2.7	723	2	T02477	hypothetical prote
787	6	2.7	724	2	A38747	phosphatidylinosit
788	6	2.7	729	2	T50989	hypothetical prote
789	6	2.7	734	2	S46765	hypothetical prote
790	6	2.7	737	2	T15597	hypothetical prote
791	6	2.7	737	2	S65758	nitrate reductase
792	6	2.7	739	2	B6434	protein F178.27 (
793	6	2.7	743	2	T26102	hypothetical prote
794	6	2.7	746	2	AC1883	nitrate reductase
795	6	2.7	749	2	S13518	transposase Tam3 -
796	6	2.7	749	2	T43370	oligonuclearyltran
797	6	2.7	749	2	S50095	splicing regulator
798	6	2.7	749	2	B87599	hypothetical prote
799	6	2.7	750	2	D86245	hypothetical prote
800	6	2.7	751	1	T29357	1-phosphatidylinos
801	6	2.7	752	2	T39338	oligosaccharyl tra
802	6	2.7	753	2	F63338	pyruvate, water di
803	6	2.7	754	1	DBBYDA	excision repair pr
804	6	2.7	754	2	S61113	Yta6 protein - yea
805	6	2.7	755	2	T20320	hypothetical prote
806	6	2.7	756	2	T50298	MAP kinase kinase
807	6	2.7	758	2	S46625	finger protein YJL
808	6	2.7	759	2	T41295	protein transport
809	6	2.7	760	2	G71417	hypothetical prote
810	6	2.7	761	2	B64449	hypothetical prote
811	6	2.7	763	2	B86454	hypothetical prote
812	6	2.7	764	2	C83513	hypothetical prote
813	6	2.7	765	2	S74598	hypothetical prote
814	6	2.7	769	2	T24949	hypothetical prote
815	6	2.7	771	2	AB7532	glucoamylase (limo
816	6	2.7	775	2	T22200	hypothetical prote
817	6	2.7	777	2	I48100	ADAM 5 protein pre
818	6	2.7	783	2	T00354	hypothetical prote
819	6	2.7	786	2	S67060	probable membrane
820	6	2.7	790	2	T24354	hypothetical prote
821	6	2.7	792	2	D64229	leucine-tRNA ligase
822	6	2.7	793	2	S73779	leucine-tRNA ligase
823	6	2.7	794	2	S55527	mammary gland fact
824	6	2.7	798	2	T15336	hypothetical prote
825	6	2.7	798	2	H87271	hypothetical prote
826	6	2.7	799	2	AD6729	TonB-dependent rec
827	6	2.7	800	2	T26683	sex-determining pr
828	6	2.7	803	2	AD1282	hypothetical prote
829	6	2.7	803	2	AH1653	leucyl-tRNA synthe
830	6	2.7	804	2	B89961	leucyl-tRNA synthe
831	6	2.7	804	2	D69650	leucine-tRNA ligase
832	6	2.7	805	2	C87861	protein P43G9.7 (I

833	6	2.7	805	2	T32377	hypothetical prote
834	6	2.7	806	2	F82899	leucyl-tRNA synth
835	6	2.7	806	2	A84060	leucyl-tRNA synth
836	6	2.7	807	2	E90523	leucyl-tRNA synth
837	6	2.7	812	2	E97831	valine-tRNA ligase
838	6	2.7	814	2	B71675	valine-tRNA ligase
839	6	2.7	814	2	T14608	hypothetical prote
840	6	2.7	815	2	B82746	DNA topoisomerase
841	6	2.7	819	2	G81698	leucyl-tRNA synth
842	6	2.7	819	2	T41544	leucine-tRNA ligas
843	6	2.7	819	2	T40527	hypothetical prote
844	6	2.7	820	2	A86510	leucyl tRNA synth
845	6	2.7	820	2	C72113	leucine-tRNA ligas
846	6	2.7	822	2	S70012	methionine-S-oxide
847	6	2.7	823	2	B87348	1,4-beta-D-glucan
848	6	2.7	824	2	T10020	leucine-tRNA ligas
849	6	2.7	826	2	T46060	hypothetical prote
850	6	2.7	826	2	T46061	hypothetical prote
851	6	2.7	828	2	AD2004	hypothetical prote
852	6	2.7	829	2	H86726	leucine-tRNA ligas
853	6	2.7	829	2	D82430	periplasmic nitrat
854	6	2.7	831	2	T05265	coat protein gamma
855	6	2.7	832	2	T49494	condensin complex
856	6	2.7	833	2	H95029	leucyl-tRNA synth
857	6	2.7	833	2	C97901	leucine-tRNA ligas
858	6	2.7	834	2	T42702	hypothetical prote
859	6	2.7	834	2	S54563	hypothetical prote
860	6	2.7	837	2	S30971	gene 26 protein -
861	6	2.7	839	2	B96538	hypothetical prote
862	6	2.7	841	2	I50428	transforming growt
863	6	2.7	845	2	AB3337	protein translocas
864	6	2.7	847	2	G75270	cation-transportin
865	6	2.7	848	1	JC1351	transforming growt
866	6	2.7	849	1	UC1350	transforming growt
867	6	2.7	851	2	T12503	hypothetical prote
868	6	2.7	851	2	T38173	probable phosphat
869	6	2.7	853	1	A41220	transforming growt
870	6	2.7	855	2	A34810	3',5'-cyclic-GMP p
871	6	2.7	861	2	T00434	probable kinesin h
872	6	2.7	862	1	FAD0A	alpha-actinin - s1
873	6	2.7	862	2	T05941	lipoxigenase (EC 1
874	6	2.7	862	2	B82312	DNA mismatch repa
875	6	2.7	863	2	A53034	gag polyprotein -
876	6	2.7	872	2	AC2134	ferrichrome-iron r
877	6	2.7	872	2	S49538	CocC protein precu
878	6	2.7	877	2	S49197	envelope protein p
879	6	2.7	883	2	S57653	brevican precursor
880	6	2.7	883	2	S49126	brevican precursor
881	6	2.7	889	2	S50934	probable membrane
882	6	2.7	891	2	T28828	hypothetical prote
883	6	2.7	892	2	T50985	related to transcr
884	6	2.7	893	2	G59431	phosphatidylinosit
885	6	2.7	894	2	T26149	hypothetical prote
886	6	2.7	895	1	PAFPA	alpha-actinin - fr
887	6	2.7	895	2	T13414	probable alpha-act
888	6	2.7	900	2	S66264	55.11 protein homo
889	6	2.7	901	1	FAHUA3	alpha-actinin 3 -
890	6	2.7	901	1	BVECCA	preprotein translo
891	6	2.7	901	2	AEO519	preprotein translo
892	6	2.7	901	2	F85492	preprotein translo
893	6	2.7	901	2	F90641	preprotein translo
894	6	2.7	904	2	A10069	preprotein translo
895	6	2.7	905	2	T00475	probable disease r
896	6	2.7	907	2	T35748	DNA polymerase I -
897	6	2.7	909	2	G69599	aconitase hydrat
898	6	2.7	910	2	A48403	alpha-actinin - Ca
899	6	2.7	912	2	A54423	brevican precursor
900	6	2.7	918	2	A88188	protein C18H9.3 [i
901	6	2.7	919	2	F83257	hypothetical prote
902	6	2.7	920	2	T26147	hypothetical prote
903	6	2.7	924	2	T13413	probable alpha-act
904	6	2.7	924	2	S06117	myosin heavy chain
905	6	2.7	925	2	T33732	probable excinucle
906	6	2.7	925	2	T00334	hypothetical prote
907	6	2.7	929	2	T38948	hypothetical colle
908	6	2.7	936	2	F75622	hypothetical prote
909	6	2.7	950	2	D70204	exinuclease ABC c
910	6	2.7	951	2	T04492	protein kinase hom
911	6	2.7	952	2	T18837	hypothetical prote
912	6	2.7	962	2	C43274	N-methyl D-asparta
913	6	2.7	966	2	T30017	hypothetical prote
914	6	2.7	968	2	D87570	exinuclease ABC.
915	6	2.7	969	2	A70912	probable leus prot
916	6	2.7	970	2	S29069	inositol polyphosp
917	6	2.7	972	2	T10023	leucine-tRNA ligas
918	6	2.7	972	2	S67048	MRP10 protein - ye
919	6	2.7	973	2	A85055	excinnase leucyl tr
920	6	2.7	974	2	AH3361	excinnase ABC c
921	6	2.7	979	2	T40006	hypothetical prote
922	6	2.7	982	1	VCLJVS	env polypotein pr
923	6	2.7	983	1	A45390	env polypotein pr
924	6	2.7	990	2	A86215	protein T6D22.8 [i
925	6	2.7	990	2	H88733	protein F32B10.3 [
926	6	2.7	1002	2	S54252	deep orange protei
927	6	2.7	1003	2	T13856	ker protein - fru
928	6	2.7	1004	2	A71617	SERA antigen/papal
929	6	2.7	1009	2	S44621	C50C3.2 protein -
930	6	2.7	1012	2	B90389	conserved hypochet
931	6	2.7	1016	2	T30990	period protein - C
932	6	2.7	1017	2	T15598	hypothetical prote
933	6	2.7	1017	2	T42384	inositol-1,4,5-tri
934	6	2.7	1018	2	T30986	period protein - C
935	6	2.7	1023	2	JC4013	major acidic nucle
936	6	2.7	1028	2	A59253	myosin I beta - hu
937	6	2.7	1028	2	S37146	myosin I heavy cha
938	6	2.7	1032	2	S53571	hypothetical prote
939	6	2.7	1043	1	SUSEMM	serine proteinase
940	6	2.7	1056	2	JQ2277	sucrose-phosphate
941	6	2.7	1061	2	A12579	conserved hypochet
942	6	2.7	1061	2	H97361	hypothetical prote
943	6	2.7	1062	2	D96540	hypothetical prote
944	6	2.7	1070	2	C75506	hypothetical prote
945	6	2.7	1071	2	S44798	F09G8.5 protein -
946	6	2.7	1071	2	S48378	probable membrane
947	6	2.7	1075	2	T48805	hypothetical prote
948	6	2.7	1078	2	T30879	dyein heavy chain
949	6	2.7	1084	2	T13173	RNA-directed DNA p
950	6	2.7	1085	2	S62516	hypothetical colle
951	6	2.7	1086	2	S74251	phosphorylase kina
952	6	2.7	1087	2	S58147	protein kinase - f
953	6	2.7	1092	2	T20117	hypothetical prote
954	6	2.7	1093	2	S74250	phosphorylase kina
955	6	2.7	1093	2	A11758	phosphorylase kina
956	6	2.7	1093	2	B40793	phosphorylase kina
957	6	2.7	1106	1	PFH0DB	platelet-derived g
958	6	2.7	1107	2	S61667	probable membrane
959	6	2.7	1118	2	S75309	hypothetical prote
960	6	2.7	1122	2	B26437	period clock prote
961	6	2.7	1127	2	A25018	circadian rhythm p
962	6	2.7	1128	2	A49660	bud emergence prot
963	6	2.7	1137	2	JC5950	integrin alpha-7 c
964	6	2.7	1151	2	T33777	hypothetical prote
965	6	2.7	1159	2	A64505	p115 homolog - Met
966	6	2.7	1173	2	T43527	sp8 protein - fibs
967	6	2.7	1173	2	T30608	proteophosphoglyca
968	6	2.7	1174	2	A40853	potassium channel
969	6	2.7	1175	2	D35815	myosin heavy chain
970	6	2.7	1176	2	C35815	myosin heavy chain
971	6	2.7	1176	2	C26427	period clock prote
972	6	2.7	1178	2	S57535	probable membrane
973	6	2.7	1178	2	S78475	mannosylphosphoryl
974	6	2.7	1186	2	S72229	meiotic recombinat
975	6	2.7	1191	2	T14154	serine/threonine p
976	6	2.7	1192	2	T17089	homeodomain-intera
977	6	2.7	1201	2	A35815	myosin heavy chain
978	6	2.7	1201	2	B35815	myosin heavy chain

979 6 2.7 1206 2 D84542 probable chloropia
980 6 2.7 1208 2 T23467 hypothetical prote
981 6 2.7 1209 2 T46027 hypothetical prote
982 6 2.7 1215 2 C84848 hypothetical prote
983 6 2.7 1218 2 A25588 period clock prote
984 6 2.7 1218 2 A26427 period clock prote
985 6 2.7 1233 2 T15316 hypothetical prote
986 6 2.7 1239 2 T49705 glutamate receptor
987 6 2.7 1240 2 S52734 hypothetical prote
988 6 2.7 1250 1 B45219 N-methyl-D-asparta
989 6 2.7 1254 2 T47141 hypothetical prote
990 6 2.7 1258 2 A12155 WD-repeat protein
991 6 2.7 1265 2 T51314 probable CO-induce
992 6 2.7 1274 2 T04018 hypothetical prote
993 6 2.7 1274 2 T19508 hypothetical prote
994 6 2.7 1283 2 S52500 oxysterol-binding
995 6 2.7 1284 2 T40578 hypothetical prote
996 6 2.7 1286 2 B71413 hypothetical prote
997 6 2.7 1289 2 T84505 calcium-dependent
998 6 2.7 1291 2 T06692 hypothetical prote
999 6 2.7 1295 2 S60179 pol. polypeptide ho
1000 6 2.7 1313 2 A48467 myosin heavy chain

ALIGNMENTS

RESULT 1

A44159
Spectrin beta-G chain - human
N:Alternate names: beta-spectrin general isoform, beta G-spectrin
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A44159
R:Hu, R.U.; Watanabe, M.; Bennett, V.
J. Biol. Chem. 267, 18715-18722, 1992
A:Title: Characterization of human brain cDNA encoding the general isoform of beta-spect
A:Reference number: A44159; MUID:92406787; PMID:1527002
A:Accession: A44159
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-2364 <HU>
A:Cross-references: UNIPROT:Q01082; UNIPARC:UPI000004EC67; GB:M96803; NID:G338442; PIDN:
A:Experimental source: brain
A>Note: sequence extracted from NCBI backbone (NCBI:P113399)
A:Gene: GDB:SPTBN1
A:Cross-references: GDB:120386; OMIM:182790
A:Map position: 2p21-2p21
C:Superfamily: spectrin beta-G chain; alpha-actinin actin-binding domain homology; pleck
C:Keywords: actin binding; cytoskeleton; duplication; heterodimer; membrane protein
F:53-271/Domain: alpha-actinin actin-binding domain homology <ACT>
F:501-412/Domain: spectrin/dystrophin repeat homology <SP1>
F:1697-1803/Domain: spectrin/dystrophin repeat homology <SP2>
F:2196-2305/Domain: pleckstrin repeat homology <PLK>

Query Match 44.1%; Score 97; DB 1; Length 2364;

Best Local Similarity 100.0%; Pred. No. 28-89; 0; Indels 0; Gaps 0;
Matches 97; Conservative 0; Mismatches 0;

QY 37 DERBAVOKKPTFTVNSHLARVSCRTDLYTDLRDGMILKLEVLSEGLPRTKGRK 96
DB 50 DERBAVOKKPTFTVNSHLARVSCRTDLYTDLRDGMILKLEVLSEGLPRTKGRK 109
QY 97 IHCLLENVDKALQFLKEORVHLENMGSHDIDVGNHRL 133
DB 110 IHCLLENVDKALQFLKEORVHLENMGSHDIDVGNHRL 146

RESULT 2
152577
beta-spectrin - mouse
C:Species: Mus sp. (mouse)

C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 13-Aug-1999
C:Accession: 152577
R:Bloom, M.L.; Birkenmeier, C.S.; Barker, J.E.
Blood 82, 2906-2914, 1993
A:Title: Complete nucleotide sequence of the murine erythroid beta-spectrin cDNA and the
A:Reference number: 152577; MUID:94033578; PMID:8219239
A:Accession: 152577
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2128 <RES>
A:Cross-references: UNIPARC:UPI000023884; GB:S66283; NID:G440899; PIDN:AA828600.1; PID:
C:Gene: Spnd-1
A:Introns: 418/3; 1742/2
C:Superfamily: spectrin beta chain; alpha-actinin actin-binding domain homology; spectri
C:Keywords: actin binding
F:53-271/Domain: alpha-actinin actin-binding domain homology <ACT>
F:1187-1892/Domain: spectrin/dystrophin repeat homology <SPH>

Query Match

Best Local Similarity 17.3%; Score 38; DB 2; Length 2128;
Best Local Similarity 100.0%; Pred. No. 1.2e-29;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 MRHCLLENVDKALQFLKEORVHLENMGSHDIDVGNHRL 132
DB 108 MRHCLLENVDKALQFLKEORVHLENMGSHDIDVGNHRL 145

RESULT 3

SCHUB
Spectrin beta chain - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A37064; J0554; A28777; A39885; B23659; B42872; B27016; A2514
R:Winkelmann, J.C.; Chang, J.G.; Tee, W.T.; Scarpa, A.L.; Marchesi, V.T.; Forget, B.G.
J. Biol. Chem. 265, 11827-11832, 1990
A:Title: Full-length sequence of the cDNA for human erythroid beta-spectrin.
A:Reference number: A37064; MUID:90307707; PMID:2195026
A:Accession: A37064
A:Molecule type: mRNA
A:Residues: 1-2137 <WIN>
A:Cross-references: UNIPROT:P11277; UNIPARC:UPI0000135DB5; GB:J05500
A>Note: the authors translated the codon GAG for residue 536 as Arg, CAC for residue 801
R:Gallagher, P.G.; Tee, W.T.; Costa, F.; Scarpa, A.; Bolvin, P.; Delaunay, J.; Forget, B.
J. Biol. Chem. 266, 15154-15159, 1991
A:Title: A splice site mutation of the beta-spectrin gene causing exon skipping in heret
A:Reference number: A39346; MUID:91332035; PMID:1840591
A:Accession: A39346
A:Molecule type: DNA
A:Residues: 2002-2137 <GAL>
A:Cross-references: UNIPARC:UPI0000173DB5; GB:J05500
R:Yoon, S.H.; Kentros, C.G.; Prchal, J.T.
Gene 91, 297-302, 1990
A:Title: Identification of an unusual deletion within homologous repeats of human reticu
A:Reference number: J0554; MUID:91007291; PMID:1976574
A:Accession: J0554
A:Molecule type: mRNA
A:Residues: 928-957, 'NY', 960-982, 'L', 984-1030, 'D', 1032-1755 <YOO>
A:Cross-references: UNIPARC:UPI0000173DB6; GB:M57948
R:Winkelmann, J.C.; Lecco, T.L.; Watkins, P.C.; Eddy, R.; Shows, T.B.; Linnenbach, A.J.;
Blood 72, 328-334, 1988
A:Title: Molecular cloning of the cDNA for human erythrocyte beta-spectrin.
A:Reference number: A90733; MUID:88269838; PMID:3390609
A:Accession: A28777
A:Molecule type: mRNA
A:Residues: 1334-1373, 'R', 1375-1432; 1909-2137 <W12>
A:Cross-references: UNIPARC:UPI000016A607; UNIPARC:UPI000016A608; GB:J05500
A>Note: authors Yoon et al. show His to be the predominant residue at position 1374 and
R:Prchal, J.T.; Morley, B.J.; Yoon, S.H.; Coetzee, T.L.; Palek, J.; Conboy, J.G.; Kan, Y.
Proc. Natl. Acad. Sci. U.S.A. 84, 7468-7472, 1987
A:Title: Isolation and characterization of cDNA clones for human erythrocyte beta-spect
A:Reference number: A39885; MUID:88041127; PMID:3478706
A:Accession: A39885

A.Molecule type: mRNA
 A.Residues: 1209-1482 <PRC>
 A.Cross-references: UNIPARC:UPI000016B05C; GB:M18054; NID:G338333; PIDN:AAA60572.1; PID:
 A.Experimental source: reticulocyte
 R.Winkelmann, J.C.; Costa, F.F.; Linzie, B.L.; Forget, B.G.
 U. Biol. Chem. 265, 20449-20454, 1990
 A.Title: Beta spectrin in human skeletal muscle. Tissue-specific differential processing
 A.Reference number: A23659; MUID:91056094; PMID:2243099
 A.Accession: B23659
 A.Molecule type: mRNA
 A.Residues: 2105-2137 <M13>
 A.Cross-references: UNIPARC:UPI0000173DB7; GB:M37885
 R.Speicher, D.W.; Weglarz, L.; Desilva, T.M.
 J. Biol. Chem. 267, 14775-14782, 1992
 A.Title: Properties of human red cell spectrin heterodimer (side-to-side) assembly and
 A.Reference number: A42872; MUID:92340516; PMID:1634521
 A.Accession: B42872
 A.Molecule type: Protein
 A.Residues: 47-56; 293-302; 1837-1846 <SP2>
 A.Cross-references: UNIPARC:UPI0000173DB8; UNIPARC:UPI0000173DB9; UNIPARC:UPI0000173DBA
 R.Speicher, D.W.; Marchesi, V.T.
 Nature 311, 177-180, 1984
 A.Title: Erythrocyte spectrin is comprised of many homologous triple helical segments.
 A.Reference number: A93341; MUID:84295638; PMID:6472478
 A.Accession: B27016
 A.Molecule type: Protein
 A.Residues: 292-324, 'X', 326-329, 'Y', 331-332; 434-532; 718-734, 'V', 736-773, 'X', 775-777; 1036
 1994-1997 <SPB>
 A.Cross-references: UNIPARC:UPI0000173DBB; UNIPARC:UPI0000173DBC; UNIPARC:UPI0000173DBD;
 A.Note: The purified protein had a blocked amino end
 C.Comment: Spectrin is a major structural component of the erythrocyte membrane cytoskel
 C.Genetics:
 A.Gene: GDB:SPTB
 A.Cross-references: GDB:119602; OMIM:182870
 A.Map position: 14q23-14q23
 C.Superfamily: spectrin beta chain; alpha-actinin actin-binding domain homology; spectri
 C.Keywords: actin binding; cytoskeleton; duplication; erythrocyte; heterodimer; membrane
 F.53-271/Domain: alpha-actinin actin-binding domain homology <ACT>
 F.301-412/Domain: spectrin/dystrophin repeat homology <SP01>
 F.421-526/Domain: spectrin/dystrophin repeat homology <SP02>
 F.527-633/Domain: spectrin/dystrophin repeat homology <SP03>
 F.636-741/Domain: spectrin/dystrophin repeat homology <SP04>
 F.742-846/Domain: spectrin/dystrophin repeat homology <SP05>
 F.847-952/Domain: spectrin/dystrophin repeat homology <SP06>
 F.953-1059/Domain: spectrin/dystrophin repeat homology <SP07>
 F.1060-1166/Domain: spectrin/dystrophin repeat homology <SP08>
 F.1167-1272/Domain: spectrin/dystrophin repeat homology <SP09>
 F.1273-1377/Domain: spectrin/dystrophin repeat homology <SP10>
 F.1378-1476/Domain: spectrin/dystrophin repeat homology <SP11>
 F.1477-1582/Domain: spectrin/dystrophin repeat homology <SP12>
 F.1583-1688/Domain: spectrin/dystrophin repeat homology <SP13>
 F.1689-1795/Domain: spectrin/dystrophin repeat homology <SP14>
 F.1796-1901/Domain: spectrin/dystrophin repeat homology <SP15>
 F.1902-2007/Domain: spectrin/dystrophin repeat homology <SP16>
 F.2008-2118/Domain: spectrin/dystrophin repeat homology <SP17>
 Query Match 17.3%; Score 38; DB 1; Length 2137;
 Best Local Similarity 100.0%; Pred. No. 1.2e-29;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 MRHCLENVKALQFLKEORVHLNMGSHDIVDGNHRL 132
 DB 108 MRHCLENVKALQFLKEORVHLNMGSHDIVDGNHRL 145
 RESULT 4
 J80271
 beta spectrin, beta SpIII signal - rat
 C.Species: Rattus norvegicus (Norway rat)
 C.Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004
 C.Accession: J80271
 R.Sakaguchi, G.; Orita, S.; Naito, A.; Maeda, M.; Igarashi, H.; Sasaki, T.; Takai, Y.
 Biochem. Biophys. Res. Commun. 248, 846-851, 1998

A>Title: A novel brain-specific isoform of beta spectrin: Isolation and its interaction
 A.Reference number: J80271; MUID:98369639; PMID:9704016
 A.Accession: J80271
 A.Molecule type: mRNA
 A.Residues: 1-238 <SAK>
 A.Cross-references: UNIPROT:Q9QW08; UNIPARC:UPI0000167B0; DDBJ:AB00134
 A.Experimental source: brain
 C.Comment: This protein is especially expressed in brain, where it was enriched in the s
 C.Superfamily: spectrin beta-G chain; alpha-actinin actin-binding domain homology <ACT>
 F.304-415/Domain: spectrin/dystrophin repeat homology <SPH>
 Query Match 12.3%; Score 27; DB 2; Length 238;
 Best Local Similarity 100.0%; Pred. No. 1.9e-18;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 87 LKPTKGMRIHCLENVKALQFLKEQ 113
 DB 103 LKPTKGMRIHCLENVKALQFLKEQ 129
 RESULT 5
 A6147
 spectrin beta chain - fruit fly (Drosophila melanogaster)
 C.Species: Drosophila melanogaster
 C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C.Accession: A6147; A33657
 R.Byers, T.J.; Branton, E.; Lue, R.A.; Winograd, E.; Branton, D.
 Proc. Natl. Acad. Sci. U.S.A. 89, 6187-6191, 1992
 A.Title: The complete sequence of Drosophila beta-spectrin reveals supra-motifs compris
 A.Reference number: A6147; MUID:92335263; PMID:1631106
 A.Accession: A6147
 A>Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-2291 <BE>
 A.Cross-references: UNIPROT:Q00963; UNIPARC:UPI000016B0B; GB:M92288; NID:G157019; PIDN:
 A.Note: sequence extracted from NCBI backbone (NCBI:108607)
 R.Byers, T.J.; Huesin-Chishti, A.; Dubreuil, R.R.; Branton, D.; Goldstein, L.S.B.
 J. Cell Biol. 109, 1633-1641, 1989
 A.Title: Sequence similarity of the amino-terminal domain of Drosophila beta spectrin to
 A.Reference number: A33657; MUID:90009037; PMID:2671025
 A.Accession: A33657
 A.Molecule type: preliminary
 A.Status: preliminary
 A.Molecule type: mRNA
 A.Residues: 1-800 <BY2>
 A.Cross-references: UNIPARC:UPI0000173DC2; GB:M92288
 C.Genetics:
 A.Gene: FlyBase:beta-Spec
 A.Cross-references: FlyBase:FBgn0003471
 C.Superfamily: spectrin beta-G chain; alpha-actinin actin-binding domain homology; plect
 C.Keywords: actin binding; cytoskeleton
 F.49-267/Domain: alpha-actinin actin-binding domain homology <ACT>
 F.297-408/Domain: spectrin/dystrophin repeat homology <SP1>
 F.417-522/Domain: spectrin/dystrophin repeat homology <SP2>
 F.1698-1804/Domain: spectrin/dystrophin repeat homology <SP3>
 F.2146-2257/Domain: spectrin/dystrophin repeat homology <PLK>
 Query Match 8.6%; Score 19; DB 1; Length 2291;
 Best Local Similarity 100.0%; Pred. No. 2.4e-10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 75 LKLELVISGERLPRPTKG 93
 DB 84 LKLELVISGERLPRPTKG 102
 RESULT 6
 S21195
 spectrin beta chain - pig
 C.Species: Sus scrofa domestica (domestic pig)
 C.Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Dec-2004
 C.Accession: S21195
 R.Frappelet, T.; Derancourt, J.; Pradel, L.A.

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #ext_change 17-May-2002
C/Accession: F59404
R/Fuchs, P.; Zorer, M.; Reznicek, G.A.; Spazlerer, D.; Oehler, S.; Castanon, M.J.; Haug
Hum. Mol. Genet. 8: 2461-2472, 1999
A>Title: Unusual 5' transcript complexity of plectin isoforms: novel tissue-specific exc
A/Reference number: F59404; MUID:20025755; PMID:10556294
A/Accession: F59404
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-214 <STO>
A/Cross-references: UNIPARC:UPI000016C902; GB:AAPI8069; NID:g6578737; PIDN:AAPI8069.1
C/Superfamily: rat ribosomal protein S10; ribosomal protein S10 homology

Query Match: 5.0%; Score 11; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTKWN 52
|||||
DB 186 VOKKFTKWN 196

RESULT 12
S66292
actin-crosslinking protein ACF7 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #ext_change 09-Jul-2004
C/Accession: S66292; S66291
R/Byers, T.J.; Beggs, A.H.; McNally, E.M.; Kunkel, L.M.
submitted to the EMBL Data Library, June 1995
A/Description: Novel actin crosslinker superfamily member identified by a two step degen
A/Reference number: S66292
A/Accession: S66292
A/Molecule type: DNA
A/Residues: 1-385 <BYE>
A/Cross-references: UNIPARC:UPI000070CP4; EMBL:L40626; NID:g845507; PID
R/Byers, T.J.; Beggs, A.H.; McNally, E.M.; Kunkel, L.M.
FBS Lett. 368, 500-504, 1995
A>Title: Novel actin crosslinker superfamily member identified by a two step degenerate
A/Reference number: S66291; MUID:95361933; PMID:7635207
A/Accession: S66291
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 97-309 <BYW>
A/Cross-references: UNIPARC:UPI0000177DB4; EMBL:L40626
C/Genetics:
A/Gene: GDB:ACF7
A/Cross-references: GDB:1323194
A/Map position: 1pter-1qter
C/Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S
F.94-308/Domain: alpha-actinin actin-binding domain homology <ACT>

Query Match: 5.0%; Score 11; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTKWN 52
|||||
DB 96 VOKKFTKWN 106

RESULT 13
D59404
plectin isoform plectin 1, alpha [imported] - mouse
C/Species: Mus musculus (house mouse)
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #ext_change 01-Feb-2002
C/Accession: D59404
R/Fuchs, P.; Zorer, M.; Reznicek, G.A.; Spazlerer, D.; Oehler, S.; Castanon, M.J.; Haug
Hum. Mol. Genet. 8: 2461-2472, 1999
A>Title: Unusual 5' transcript complexity of plectin isoforms: novel tissue-specific exc
A/Reference number: D59404; MUID:20025755; PMID:10556294
A/Accession: D59404
A/Status: preliminary

A/Molecule type: DNA
A/Residues: 1-964 <STO>
A/Cross-references: UNIPARC:UPI0000021F4E; GB:AAPI8068; NID:g6578735; PIDN:AAPI8068.1

Query Match: 5.0%; Score 11; DB 2; Length 964;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTKWN 52
|||||
DB 186 VOKKFTKWN 196

RESULT 14
T22552
hypothetical protein ZK1151.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 29-Oct-1999
C/Accession: T22552; T27703
R/Harris, B.
submitted to the EMBL Data Library, March 1997
A/Reference number: Z19580
A/Accession: T22552
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1355 <WIL>
A/Cross-references: UNIPARC:UPI000017BCFE; EMBL:Z92788; PIDN:CAB07214.1; GSPDB:GN00019;
A/Experimental source: clone F51B8
R/Harris, B.
submitted to the EMBL Data Library, March 1997
A/Reference number: Z20408
A/Accession: T27703
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1355 <W12>
A/Cross-references: UNIPARC:UPI000017BCEF; EMBL:Z93398; PIDN:CAB07724.1; GSPDB:GN00019;
A/Experimental source: clone ZK1151
C/Genetics:
A/Gene: CESP:ZK1151.1
A/Map position: 1
A/Introns: 94/3; 124/3; 150/3; 209/3; 248/3; 311/3; 424/3; 734/2; 934/2; 1104/3; 1167/3;

Query Match: 5.0%; Score 11; DB 2; Length 1355;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTKWN 52
|||||
DB 79 VOKKFTKWN 89

RESULT 15
T42725
actin binding protein ACF7, neural isoform 1 - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #ext_change 03-Nov-2000
C/Accession: T42725
R/Bernier, G.; Machieu, M.; De Repentigny, Y.; Vidal, S.M.; Kothary, R.
Genomics 38, 19-29, 1996
A>Title: Cloning and characterization of mouse ACF7, a novel member of the dystronin subf.
A/Reference number: Z20900; MUID:97124842; PMID:8934775
A/Accession: T42725
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1825 <EBR>
A/Cross-references: UNIPARC:UPI000016D118; EMBL:U67203; NID:g1675221; PID:g1675222; PIDN
C/Genetics:
A/Gene: ACF7
A/Map position: 4
C/Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S
C/Keywords: actin binding

Query Match: 5.0%; Score 11; DB 2; Length 1825;

A:Cross-references: UNIPARC:UPI0000202A1A9; EMBL:AJ011925; NID:G3758910; PIDD:CAA09870.1;
C:Genetics:
A:Gene: kak
A:Cross-references: FlyBase:FBgn0013733
A:Note: kak

Query Match 4.5%; Score 10; DB 2; Length 2396;
Best Local Similarity 100.0%; Pred. No. 0.33;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 OKKFTTKWN 52
|||||
DB 150 OKKFTTKWN 159

RESULT 20
S06461
dyctrophin, brain - rat (fragment)
N:Alternate names: Duchenne muscular dystrophy protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 09-Jul-2004
C:Accession: S06461
R:Undel, U.; Zuk, D.; Binat, P.; Zeelson, E.; Levy, Z.; Neuman, S.; Yaffe, D.
Mature 337, 76-78, 1989
A:Title: Duchenne muscular dystrophy gene product is not identical in muscle and brain.
A:Reference number: S06461; MUID:89082658; PMID:2909892
A:Accession: S06461
A:Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-98 <RES>
A:Cross-references: UNIPROT:Q63770; UNIPARC:UPI000017768B; EMBL:X14182
C:Genetics:
A:Introns: 3/1
A:Note: the list of introns may be incomplete
C:Superfamily: dyctrophin; alpha-actinin actin-binding domain homology; spectrin/dystroph
C:Keyword: alternative splicing
F:3-98/Domain: alpha-actinin actin-binding domain homology (fragment) <AAH>

Query Match 4.1%; Score 9; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTTKW 50
|||||
DB 8 VOKKFTTKW 16

RESULT 21
I39161
dyctonin isoform 2 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 17-Nov-2000
C:Accession: I39161
R:Brown, A.; Dalpe, G.; Mathieu, M.; Kothary, R.
Genomics 29, 777-780, 1995
A:Title: Cloning and characterization of the neural isoforms of human dyctonin.
A:Reference number: I39160; MUID:96121394; PMID:8515775
A:Accession: I39161
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-230 <RES>
A:Cross-references: UNIPARC:UPI000016A21F; EMBL:U31851; NID:G1049105; PIDD:AAC50244.1; F
C:Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S

Query Match 4.1%; Score 9; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTTKW 50
|||||
DB 220 VOKKFTTKW 228

RESULT 22
J80233
tropoinin-I - scallop (Chlamys nipponensis)

C:Species: Chlamys nipponensis (Japanese scallop)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 09-Jul-2004
C:Accession: J80233
R:Tanaka, H.; Ojima, T.; Nishita, K.
J. Biochem. 124, 304-310, 1998
A:Title: Amino acid sequence of tropoinin-I from Akazara scallop striated adductor muscle
A:Reference number: J80233; MUID:98351986; PMID:9685719
A:Accession: J80233
A:Molecule type: protein
A:Residues: 1-292 <TAB>
A:Cross-references: UNIPROT:Q7M3Y3; UNIPARC:UPI0000177780
C:Superfamily: tropoinin I

Query Match 4.1%; Score 9; DB 2; Length 292;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 RROQREER 149
|||||
DB 104 RROQREER 112

RESULT 23
I39160
dyctonin isoform 1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 17-Nov-2000
C:Accession: I39160
R:Brown, A.; Dalpe, G.; Mathieu, M.; Kothary, R.
Genomics 29, 777-780, 1995
A:Title: Cloning and characterization of the neural isoforms of human dyctonin.
A:Reference number: I39160; MUID:96121394; PMID:8515775
A:Accession: I39160
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-461 <RES>
A:Cross-references: UNIPARC:UPI000013CB6B; EMBL:U31850; NID:G1049103; PIDD:AAC50243.1; P
C:Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S
F:34-248/Domain: alpha-actinin actin-binding domain homology <ACT>

Query Match 4.1%; Score 9; DB 2; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTTKW 50
|||||
DB 36 VOKKFTTKW 44

RESULT 24
I49290
dyctonin - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I49290
R:Brown, A.; Bernier, G.; Mathieu, M.; Rosant, J.; Kothary, R.
Nature Genet. 10, 301-306, 1995
A:Title: The mouse dyctonin muscular gene is a neural isoform of bullous pemphigoid an
A:Reference number: I49290; MUID:95400297; PMID:7670468
A:Accession: I49290
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-511 <RES>
A:Cross-references: UNIPROT:Q60824; UNIPARC:UPI000016CB1A; EMBL:U22452; NID:G904018; P
C:Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S
F:34-248/Domain: alpha-actinin actin-binding domain homology <ACT>

Query Match 4.1%; Score 9; DB 2; Length 511;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKTFTKW 50
|||||
DB 36 VOKKTFTKW 44

RESULT 25

C75120

hypothetical protein PAB1857 - *Pyrococcus abyssi* (strain Orsay)C.Species: *Pyrococcus abyssi*

C.Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C.Accession: C75120

R.anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A.Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome stru

A.Reference number: A75001

A.Accession: C75120

A.Status: preliminary

A.Molecule type: DNA

A.Residues: 1-602 <KAW>

A.Cross-references: UNIPROT:Q9VOM7; UNIPARC:UPI000006331C; GB:AJ248285; GB:AL096836; NID

A.Experimental source: strain Orsay

C.Genetics:

A.Gene: PAB1857

C.Superfamily: *Pyrococcus horikoshii* hypothetical protein PH1386

Query Match

Best Local Similarity 4.1%; Score 9; DB 2; Length 602;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 GRPKQLQDE 38
|||||
DB 126 GRPKQLQDE 134

RESULT 26

I49298

dystonin isoform 2 - mouse (fragment)

C.Species: *Mus musculus* (house mouse)

C.Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 17-Nov-2000

C.Accession: I49298

R.Brown, A.; Beninier, G.; Mathieu, M.; Rossant, J.; Kothary, R.

Nature Genet. 10, 301-306, 1995

A.Title: The mouse dystonia musculorum gene is a neural isoform of bullous pemphigoid an

A.Reference number: I49290; MUID:95400297; PMID:7670468

A.Accession: I49298

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-686 <RES>

A.Cross-references: UNIPARC:UPI0000028749; EMBL:U25158; NID:G904021; PIDN:AAC52231.1; PI

C.Superfamily: plectin; alpha-actinin actin-binding domain homology; ribosomal protein S

F.209-423/Domain: alpha-actinin actin-binding domain homology <ACT>

Query Match

Best Local Similarity 4.1%; Score 9; DB 2; Length 686;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKTFTKW 50
|||||
DB 211 VOKKTFTKW 219

RESULT 27

I15570

hypothetical protein C23F12.2 - *Caenorhabditis elegans*C.Species: *Caenorhabditis elegans*

C.Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C.Accession: I15570

R.Fulton, B.

submitted to the EMBL Data Library, October 1995

A.Description: The sequence of C. elegans cosmid C23F12.

A.Reference number: Z18371

A.Accession: I15570

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-874 <Ful>

A.Cross-references: UNIPROT:Q18109; UNIPARC:UPI000007B862; EMBL:U39676; NID:G1049427; P

A.Experimental source: strain Bristol N2; clone C23F12

C.Genetics:

A.Gene: CESP:C23F12.2

A.Map position: X

A.Introns: 27/3; 136/3; 240/3; 292/1; 319/1; 346/3; 368/1; 405/3; 421/3; 496/3; 525/3;

Query Match

Best Local Similarity 4.1%; Score 9; DB 2; Length 874;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GSHDIVDGN 129
|||||
DB 94 GSHDIVDGN 102

RESULT 28

S28381

utrophin - human

N.Alternate names: dystrophin-related protein

C.Species: *Homo sapiens* (man)

C.Date: 17-Apr-1993 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004

C.Accession: S28381; S28914; S03966

R.Finsley, J.M.

submitted to the EMBL Data Library, November 1992

A.Reference number: S28381

A.Accession: S28381

A.Molecule type: mRNA

A.Residues: 1-3433 <TINI>

A.Cross-references: UNIPROT:P46939; UNIPARC:UPI00001376EA; EMBL:X69086; NID:G34811; PIDN

R.Finsley, J.M.; Blake, D.J.; Roche, A.; Fairbrother, U.; Riss, J.; Byth, B.C.; Knight,

Nature 360, 591-593, 1992

A.Title: Primary structure of dystrophin-related protein.

A.Reference number: S28914; MUID:9306045; PMID:1461283

A.Accession: S28914

A.Molecule type: mRNA

A.Residues: 27-246; 2839-3443 <TIN2>

A.Cross-references: UNIPARC:UPI000173E73; UNIPARC:UPI000173E74; EMBL:X69086

A.Love, D.R.; Hill, D.F.; Dickson, G.; Spurr, N.K.; Byth, B.C.; Marsden, R.F.; Walsh, F.

Nature 339, 55-58, 1989

A.Title: An autosomal transcript in skeletal muscle with homology to dystrophin.

A.Reference number: S03966; MUID:89238543; PMID:2541343

A.Accession: S03966

A.Molecule type: mRNA

A.Residues: 2944-3433 <LOV>

A.Cross-references: UNIPARC:UPI000000673; EMBL:X15488; NID:G30933; PIDN:CAA33515.1; PID

C.Comment: This protein is found primarily at the neuromuscular junctions in adult muscul

d regenerating muscle.

C.Genetics:

A.Gene: GDB:UTRN; DMDL

A.Cross-references: GDB:119851; OMIM:128240

A.Map position: 6q24-q24

C.Superfamily: dystrophin; alpha-actinin actin-binding domain homology; spectrin/dystrop

C.Keywords: actin binding; cytoskeleton; leucine zipper; membrane-associated protein; mu

F.30-248/Domain: alpha-actinin actin-binding domain homology <ACT>

F.308-417/Domain: spectrin/dystrophin repeat homology <SP01>

F.418-526/Domain: spectrin/dystrophin repeat homology <SP02>

F.528-637/Domain: spectrin/dystrophin repeat homology <SP03>

F.638-685/Region: hinge

F.686-796/Domain: spectrin/dystrophin repeat homology <SP04>

F.804-902/Domain: spectrin/dystrophin repeat homology <SP05>

F.906-1013/Domain: spectrin/dystrophin repeat homology <SP06>

F.1015-1122/Domain: spectrin/dystrophin repeat homology <SP07>

F.1124-1230/Domain: spectrin/dystrophin repeat homology <SP08>

F.1233-1334/Domain: spectrin/dystrophin repeat homology <SP09>

F.1339-1450/Domain: spectrin/dystrophin repeat homology <SP10>

F.1451-1541/Domain: spectrin/dystrophin repeat homology #status atypical <SP11>

F.1543-1649/Domain: spectrin/dystrophin repeat homology <SP12>

F.1651-1755/Domain: spectrin/dystrophin repeat homology <SP13>

F.1856-1973/Domain: spectrin/dystrophin repeat homology <SP14>

F:1975-2081/Domain: spectrin/dystrophin repeat homology <SP15>
 F:2083-2185/Domain: spectrin/dystrophin repeat homology <SP16>
 F:2227-2333/Domain: spectrin/dystrophin repeat homology <SP17>
 F:2333-2440/Domain: spectrin/dystrophin repeat homology <SP18>
 F:2444-2556/Domain: spectrin/dystrophin repeat homology <SP19>
 F:2556-2688/Domain: spectrin/dystrophin repeat homology <SP20>
 F:2690-2797/Domain: spectrin/dystrophin repeat homology <SP21>
 F:2798-2869/Domain: spectrin/dystrophin repeat homology <SP22>
 F:2812-2849/Domain: spectrin/dystrophin repeat homology <SP23>
 F:2837-3117/Region: spectrin/dystrophin repeat homology <SP24>
 F:3263-3284/Region: spectrin/dystrophin repeat homology <SP25>
 F:3328-3349/Region: spectrin/dystrophin repeat homology <SP26>

Query Match 4.1% Score 9; DB 1; Length 3433;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTTKW 50
 DB 32 VOKKFTTKW 40

RESULT 29
 S02041
 Dystrophin, muscle - chicken
 N:Alternate names: duchenne muscular dystrophy protein
 C:Species: Gallus gallus (chicken)
 C>Date: 07-Sep-1990 #sequence revision 27-Jun-1994 #text_change 09-Jul-2004
 C/Accession: S02041; S02013; S71487
 R:Lemaire, C.; Heilig, R.; Mandel, J.L.
 Nucleic Acids Res. 16, 11815-11816, 1988
 A>Title: Nucleotide sequence of chicken dystrophin cDNA.
 A/Reference number: S02041; PMID:89098331; PMID:3062582
 A/Accession: S02041
 A/Status: translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-3660 <LEM>
 A/Cross-references: UNIPROT:P11533; UNIPARC:UPI00001294D5; EMBL:X13369; NID:963369; PIDN:
 A/Note: 1869-His, 1885-Arg, and sequences lacking 1171-Met were also found
 R:Lemaire, C.; Heilig, R.; Mandel, J.L.
 EMBO J. 7, 4157-4162, 1988
 A>Title: The chicken dystrophin cDNA: striking conservation of the C-terminal coding and
 A/Reference number: S02013; PMID:92210800; PMID:3072195
 A/Accession: S02013
 A/Status: nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-3573, HA, 3576-3660 <LEM2>
 A/Cross-references: UNIPARC:UPI0000173E71
 R:Heilig, R.; Lemaire, C.; Mandel, J.L.
 Nucleic Acids Res. 15, 9129-9142, 1987
 A>Title: A 230kb cosmid walk in the Duchenne muscular dystrophy gene: detection of a con
 A/Reference number: S09071; PMID:86067745; PMID:2825128
 A/Accession: S71487
 A/Molecule type: DNA
 A/Residues: 222-281 <HE1>
 A/Cross-references: UNIPARC:UPI0000173E72
 C/Comment: Dystrophin is proposed to play a role in anchoring the cytoskeleton to the pl
 C/Superfamily: spectrin/dystrophin repeat homology <SP05>
 C/Keywords: actin binding; calmodulin binding; cytoskeleton; leucine zipper; membrane-ae
 F:18-237/Domain: alpha-actinin actin-binding domain homology <ACT>
 F:253-327/Region: hinge
 F:340-449/Domain: spectrin/dystrophin repeat homology <SP01>
 F:450-558/Domain: spectrin/dystrophin repeat homology <SP02>
 F:560-669/Domain: spectrin/dystrophin repeat homology <SP03>
 F:670-719/Region: hinge
 F:720-830/Domain: spectrin/dystrophin repeat homology <SP04>
 F:838-936/Domain: spectrin/dystrophin repeat homology <SP05>
 F:940-1047/Domain: spectrin/dystrophin repeat homology <SP06>
 F:1049-1156/Domain: spectrin/dystrophin repeat homology <SP07>
 F:1158-1265/Domain: spectrin/dystrophin repeat homology <SP08>
 F:1267-1369/Domain: spectrin/dystrophin repeat homology <SP09>
 F:1374-1479/Domain: spectrin/dystrophin repeat homology <SP10>

F:1460-1570/Domain: spectrin/dystrophin repeat homology #status atypical <SP11>
 F:1572-1678/Domain: spectrin/dystrophin repeat homology <SP12>
 F:1680-1784/Domain: spectrin/dystrophin repeat homology <SP13>
 F:1787-1877/Domain: spectrin/dystrophin repeat homology #status atypical <SP14>
 F:1878-1984/Domain: spectrin/dystrophin repeat homology <SP15>
 F:1986-2103/Domain: spectrin/dystrophin repeat homology <SP16>
 F:2105-2211/Domain: spectrin/dystrophin repeat homology <SP17>
 F:2213-2319/Domain: spectrin/dystrophin repeat homology <SP18>
 F:2323-2419/Domain: spectrin/dystrophin repeat homology <SP19>
 F:2420-2467/Region: hinge
 F:2468-2574/Domain: spectrin/dystrophin repeat homology <SP20>
 F:2576-2683/Domain: spectrin/dystrophin repeat homology <SP21>
 F:2685-2799/Domain: spectrin/dystrophin repeat homology <SP22>
 F:2801-2928/Domain: spectrin/dystrophin repeat homology <SP23>
 F:2930-3037/Domain: spectrin/dystrophin repeat homology <SP24>
 F:3038-3075/Region: hinge
 F:3079-3089/Domain: spectrin/dystrophin repeat homology <SP25>
 F:3079-3357/Region: spectrin/dystrophin repeat homology <SP26>
 F:3481-3502/Region: spectrin/dystrophin repeat homology <SP27>
 F:3547-3568/Region: spectrin/dystrophin repeat homology <SP28>

Query Match 4.1% Score 9; DB 1; Length 3660;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VOKKFTTKW 50
 DB 20 VOKKFTTKW 28

RESULT 30
 S28916
 Dystrophin - mouse
 N:Alternate names: duchenne muscular dystrophy protein
 C:Species: Mus musculus (house mouse)
 C>Date: 22-Nov-1993 #sequence revision 03-Nov-1995 #text_change 09-Jul-2004
 C/Accession: S28916; B27162; S10922; C43837; B40134
 R:Giles, R.D.; Phelps, S.F.; Cortez, M.D.; Roberts, R.; Caskey, C.T.; Chamberlain, J.S.
 Nucleic Acids Res. 20, 1725-1731, 1992
 A>Title: Human and murine dystrophin mRNA transcripts are differentially expressed durin
 A/Reference number: S28916; PMID:92253376; PMID:1579466
 A/Accession: S28916
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-3678 <BIE>
 A/Cross-references: UNIPROT:P11531; UNIPARC:UPI0000279E7; EMBL:M68859
 A/Note: The nucleotide sequence was submitted to the EMBL Data Library, October 1991
 R:Koenig, M.; Hoffman, B.P.; Bertelson, C.J.; Monaco, A.P.; Feener, C.; Kunkel, L.M.
 Cell 50, 509-517, 1987
 A>Title: Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary
 A/Reference number: A90897; PMID:87273512; PMID:3607877
 A/Accession: B27162
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-201 <MOE>
 A/Cross-references: UNIPARC:UPI0000177666
 R:Nudel, U.; Zuk, D.; Binat, P.; Zeelon, E.; Levy, Z.; Neuman, S.; Yaffe, D.
 Nature 337, 76-78, 1989
 A>Title: Duchenne muscular dystrophy gene product is not identical in muscle and brain.
 A/Reference number: S06461; PMID:89082658; PMID:2909692
 A/Accession: S10922
 A/Status: translation not shown
 A/Molecule type: mRNA
 A/Residues: 1-106 <NUD>
 A/Cross-references: UNIPARC:UPI0000177687; EMBL:X14183
 R:Rapaport, D.; Lederfein, D.; den Dunnen, J.T.; Grootscholten, P.M.; Van Ommen, G.J.;
 Differentiation 49, 187-193, 1992
 A>Title: Characterization and cell type distribution of a novel, major transcript of th
 A/Reference number: A43837; PMID:92316332; PMID:1377655
 A/Accession: C43837
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 3069-3181 <RAP>

A:Cross-references: UNIPARC:UPI0000177668
 A:Note: sequence extracted from NCBI backbone
 R:Hoffman, E.P.; Monaco, A.P.; Feener, C.C.; Kunkel, L.M.
 Science 238, 347-350, 1987
 A:Title: Conservation of the Duchenne muscular dystrophy gene in mice and humans.
 A:Reference number: A40134; MUID:88018015; PMID:3659917
 A:Accession: B40134
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 300-676; 'F', 678-1390 <HOF>
 A:Cross-references: UNIPARC:UPI0000177689; GB:M18025
 C:Genetics:
 A:Introns: 11/1
 A:Note: the list of introns may be incomplete
 C:Superfamily: dystrophin; alpha-actinin actin-binding domain homology; spectrin/dystrophin
 C:Keywords: actin binding; alternative splicing
 F:14-233/Domain: alpha-actinin actin-binding domain homology <ACT>
 F:340-449/Domain: spectrin/dystrophin repeat homology <SP1>
 F:450-558/Domain: spectrin/dystrophin repeat homology <SP2>
 F:2797-2924/Domain: spectrin/dystrophin repeat homology <SP3>
 F:3048-3085/Domain: WW repeat homology <WW1>

Query Match 4.1%; Score 9; DB 2; Length 3678;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 42 YKKKFTKW 50
 DB 16 YKKKFTKW 24

RESULT 31
 F86242
 Unknown protein, 98896-95855 (imported) - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: F86242
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Holt, D.; Chang, M.K.; Con, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizart, L.
 Nature 408, 816-820, 2000
 A:Author: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Malt, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F86242
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-582 <STO>
 A:Cross-references: UNIPROT:O04096; UNIPARC:UPI00000A1725; GB:AE005172; NID:G1931653; PI
 C:Genetics:
 A:Map position: 1

Query Match 3.6%; Score 8; DB 2; Length 592;
 Best Local Similarity 100.0%; Pred. No. 9.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 145 EEEERKR 152
 DB 66 EEEERKR 73

RESULT 32
 FAHUA2
 alpha-actinin 2 - human
 C:Species: Homo sapiens (man)
 C:Date: 16-Oct-1992 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
 C:Accession: A40199
 R:Beegs, A.H.; Byers, T.J.; Knoll, J.H.M.; Boyce, F.M.; Bruns, G.A.P.; Kunkel, L.M.
 J. Biol. Chem. 267, 9281-9288, 1992

A:Title: Cloning and characterization of two human skeletal muscle alpha-actinin genes
 A:Reference number: A40199; MUID:92250531; PMID:1339456
 A:Accession: A40199
 A:Molecule type: mRNA
 A:Residues: 1-894 <BEG>
 A:Cross-references: UNIPROT:P35609; UNIPARC:UPI0000125088; GB:M86406; NID:G178053; PIDN
 C:Comment: The EF hand structures are predicted to be incapable of binding calcium.
 C:Comment: The amino-terminal domain is involved in actin binding, the middle domain in
 ng by the adjacent chain.
 C:Genetics:
 A:Gene: GDB:ACTN2
 A:Cross-references: GDB:127919; OMIM:102573
 A:Map position: 1q42-1q43
 C:Complex: homodimer
 C:Superfamily: alpha-actinin; alpha-actinin actin-binding domain homology; calmodulin
 C:Keywords: actin binding; cardiac muscle; duplication; EF hand; heart; homodimer; skele
 F:37-250/Domain: alpha-actinin actin-binding domain homology <ACT>
 F:280-391/Domain: spectrin/dystrophin repeat homology <SP1>
 F:400-506/Domain: spectrin/dystrophin repeat homology <SP2>
 F:515-627/Domain: spectrin/dystrophin repeat homology <SP3>
 F:636-740/Domain: spectrin/dystrophin repeat homology <SP4>
 F:753-785/Domain: calmodulin repeat homology <EF1>
 F:789-821/Domain: calmodulin repeat homology <EF2>

Query Match 3.6%; Score 8; DB 1; Length 894;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 83 SGERLPKP 90
 DB 80 SGERLPKP 87

RESULT 33
 S02032
 alpha-actinin 2, skeletal muscle splice form SK - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
 C:Accession: S02032; 150604; S36393; S27391; S15481
 R:Arimura, C.; Suzuki, T.; Yanagisawa, M.; Imamura, M.; Hamada, Y.; Masaki, T.
 Eur. J. Biochem. 177, 649-655, 1988
 A:Title: Primary structure of chicken skeletal muscle and fibroblast alpha-actinins deduced
 A:Reference number: S02032; MUID:89664821; PMID:3197725
 A:Accession: S02032
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-897 <ARI>
 A:Cross-references: UNIPROT:P20111; UNIPARC:UPI0000125086; GB:X13874; NID:G63787; PIDN:C
 A:Note: part of this sequence was confirmed by protein sequencing
 R:Tokune, Y.; Goto, S.; Imamura, M.; Odinata, T.; Masaki, T.; Endo, T.
 Exp. Cell Res. 197, 158-167, 1991
 A:Title: Transfection of chicken skeletal muscle alpha-actinin cDNA into nonmuscle and m
 A:Reference number: 150604; MUID:92070385; PMID:1720388
 A:Accession: 150604
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-897 <TO2>
 A:Cross-references: UNIPARC:UPI0000125086; EMBL:X59247; NID:G63030; PIDN:CAA41935.1; PID
 R:Part, T.
 submitted to the EMBL Data Library, October 1992
 A:Reference number: S36393
 A:Accession: S36393
 A:Molecule type: DNA
 A:Residues: 662-867 <PAR1>
 A:Cross-references: UNIPARC:UPI0000177553; EMBL:X68801
 R:Part, T.; Walters, G.T.; Patel, B.; Milla, D.B.; Critchley, D.R.
 Eur. J. Biochem. 210, 801-809, 1992
 A:Title: A chick skeletal muscle alpha-actinin gene gives rise to two alternatively spliced
 A:Reference number: S27391; MUID:93130905; PMID:1483465
 A:Accession: S27391
 A:Molecule type: DNA
 A:Residues: 707-867 <PAR2>
 A:Cross-references: UNIPARC:UPI0000177554; EMBL:X68801

A/Note: the authors translated the codon AAA for residue 714 as Leu and AAG for residue 715. The EF hand structures are predicted to be incapable of binding calcium.
C/Comment: The amino-terminal domain is involved in actin binding, the middle domain in actin chain.

C/Genetic: 721/3; 770/3; 792/3; 845/3
A/Introns: 1
A/Note: the list of introns is incomplete

C/Complex: homodimer
C/Superfamily: alpha-actinin; alpha-actinin actin-binding domain homology; calmodulin repeat
C/Keywords: actin binding; alternative splicing; duplication; EF hand; homodimer; skeletal

F/0-253/Domain: alpha-actinin actin-binding domain homology <ACT>
F/83-394/Domain: spectrin/dystrophin repeat homology <SP1>
F/403-509/Domain: spectrin/dystrophin repeat homology <SP2>
F/518-630/Domain: spectrin/dystrophin repeat homology <SP3>
F/639-743/Domain: spectrin/dystrophin repeat homology <SP4>
F/756-788/Domain: calmodulin repeat homology <EF1>
F/792-824/Domain: calmodulin repeat homology <EF2>

Query Match 3.6%; Score 8; DB 2; Length 897;
Best Local Similarity 100.0%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 SGERLPKP 90
|||
Db 83 SGERLPKP 90

RESULT 34

S45673
Alpha-actinin, 115K nonmuscle isoform - chicken

C/Species: Gallus gallus (chicken)

C/Date: 10-Dec-1994 #sequence_revise 03-Nov-1995 #text_change 09-Jul-2004

C/Accession: S45673; A45090

R/Imamura, M.; Sakurai, T.; Ogawa, Y.; Ishikawa, T.; Goto, K.; Masaki, T.

Bur. J. Biochem. 223, 395-401, 1994

A/Title: Molecular cloning of low-Ca(2+)-sensitive-type non-muscle alpha-actinin.

A/Reference number: S45673; MUID:9433327; PMID:8055908

A/Accession: S45673

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-904 <IMA>

A/Cross-references: UNIPROT:Q90734; UNIPARC:UPI0000125091; GB:D26597; NID:G517084; PIDN:

J. Biol. Chem. 267, 25927-25933, 1992

A/Title: A novel nonmuscle alpha-actinin. Purification and characterization of chicken

A/Reference number: A45090; MUID:93100311; PMID:1334489

A/Accession: A45090

A/Status: preliminary

A/Molecule type: protein

A/Residues: 294-316 <IM2>

A/Cross-references: UNIPARC:UPI00000PBBF

A/Experimental source: lung

A/Note: sequence extracted from NCBI backbone (NCBIP:120852)

C/Superfamily: alpha-actinin; alpha-actinin actin-binding domain homology; calmodulin re

C/Keywords: actin binding; EF hand

F/42-255/Domain: alpha-actinin actin-binding domain homology <ACT>

F/285-396/Domain: spectrin/dystrophin repeat homology <SP1>

F/403-511/Domain: spectrin/dystrophin repeat homology <SP2>

F/520-632/Domain: spectrin/dystrophin repeat homology <SP3>

F/611-745/Domain: spectrin/dystrophin repeat homology <SP4>

F/758-790/Domain: calmodulin repeat homology <EF1>

F/799-831/Domain: calmodulin repeat homology <EF2>

Query Match 3.6%; Score 8; DB 2; Length 904;
Best Local Similarity 100.0%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 SGERLPKP 90
|||
Db 83 SGERLPKP 92

RESULT 35

JC7186

alpha-actinin-4 - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 04-Mar-2000 #sequence_revise 04-Mar-2000 #text_change 09-Jul-2004

C/Accession: JC7186

R/Bl-Husseini, A.E.D.; Kwasnicka, D.; Yamada, T.; Hirohashi, S.; Vincent, S.R.

Biochem. Biophys. Res. Commun. 267, 906-911, 2000

A/Title: BERP, a novel ring finger protein, binds to alpha-actinin-4.

A/Reference number: JC7186; MUID:20139748; PMID:10673389

A/Accession: JC7186

A/Molecule type: mRNA

A/Residues: 1-911 <EIA>

A/Cross-references: UNIPROT:Q9QXQ0; UNIPARC:UPI0000125094; GB:AF190909; NID:G6636118; PI

A/Experimental source: brain

C/Comment: This protein, a calcium-insensitive and nonmuscle form of alpha-actinin impli

finger protein, that acts to anchor microfilaments to various cellular structures.

C/Superfamily: alpha-actinin; alpha-actinin actin-binding domain homology; calmodulin re

C/Keywords: brain; EF hand; microfilament

Query Match 3.6%; Score 8; DB 2; Length 911;
Best Local Similarity 100.0%; Pred. No. 14;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 SGERLPKP 90
|||
Db 92 SGERLPKP 99

RESULT 36

T36663
protein kinase, transmembrane - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 03-Dec-1999 #sequence_revise 03-Dec-1999 #text_change 09-Jul-2004

C/Accession: T36663

R/Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrett, B.G.; Rajandream, M.A

submitted to the EMBL Data Library, March 1999

A/Reference number: T21611

A/Accession: T36663

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1089 <SER>

A/Cross-references: UNIPROT:Q9X908; UNIPARC:UPI00000DAF05; EMBL:AL035636; PIDN:CA830479

A/Experimental source: strain A3(2)

C/Genetics:

A/Gene: SC0DB:SC5.05C

Query Match 3.6%; Score 8; DB 2; Length 1089;
Best Local Similarity 100.0%; Pred. No. 17;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 EBERRKR 152
|||
Db 761 EBERRKR 768

RESULT 37

E84193
chromosome segregation [imported] - Halobacterium sp. NRC-1

C/Species: Halobacterium sp. NRC-1

C/Date: 02-Feb-2001 #sequence_revise 02-Feb-2001 #text_change 09-Jul-2004

C/Accession: E84193

R/M, W.V.; Kennedy, S.P.; Mahapatra, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laaky,

Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabb,

Jung, K.H.; Alam, M.; Freilich, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; L

A/Title: Genome sequence of Halobacterium species NRC-1.

A/Reference number: A84160; MUID:20504483; PMID:11016950

A/Accession: E84193

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-1190 <STO>

A/Cross-references: UNIPROT:Q9HS95; UNIPARC:UPI00000635B5; GB:AE004437; NID:G10579965;

C/Genetics:
A/Gene: smc1

Query Match
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 35 QKTTFTKW 42
|||
Db 768 QKTTFTKW 775

RESULT 38

A37792

spectrin beta-H chain - fruit fly (*Drosophila melanogaster*) (fragment)C/Species: *Drosophila melanogaster*

C/Date: 30-Apr-1991 #sequence_revision 08-Nov-1996 #text_change 09-Jul-2004

C/Accession: A37792; S70848; S15666

R/Lundblad, R.R.; Byers, T.J.; Stewart, C.T.; Kiehart, D.P.

J. Cell Biol. 111, 1849-1858, 1990

A/Title: A beta-spectrin isoform from *Drosophila* (beta-H) is similar in size to vertebrate

A/Reference number: A37792; MUID:91035599; PMID:2229176

A/Accession: A37792

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-1645 <DUB>

A/Cross-references: UNIPROT:Q99140; UNIPARC:UPI000017BD9; GB:X53992

A/Note: Met-14 is the probable initiator

R/Lundblad, R.R.

submitted to the EMBL Data Library, July 1990

A/Accession: S70848

A/Molecule type: mRNA

A/Residues: 1-400,403-410, 'QV',411-1645 <DUB>

A/Cross-references: UNIPARC:UPI00008357B; EMBL:X53992; NID:97654; PID:97655

A/Note: Met-14 is the probable initiator

C/Genetics:

A/Gene: FlyBase:Kst

A/Cross-references: FlyBase:FBgn0004167

C/Keywords: actin binding

F/36-258/Domain: alpha-actinin actin-binding domain homology <ACT>

F/288-339/Domain: spectrin/dystrophin repeat homology <SP1>

F/400-510/Domain: spectrin/dystrophin repeat homology <SP2>

F/851-896/Domain: SH3 homology <SH3>

F/999-1091/Domain: spectrin/dystrophin repeat homology <SP3>

F/1303-1408/Domain: spectrin/dystrophin repeat homology <SP4>

F/1409-1512/Domain: spectrin/dystrophin repeat homology <SP5>

Query Match

Best Local Similarity 100.0%; Pred. No. 25;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 QKTTFTKW 50
|||
Db 39 QKTTFTKW 46

RESULT 39

T42047

insulin receptor homolog - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C/Accession: T42047

R/Knutz, K.D.; Tissenbaum, H.A.; Liu, Y.; Ruvkun, G.

Science 277, 942-945, 1997

A/Title: Daf-2, an insulin receptor-like gene that regulates longevity and diapause in *C.*

A/Reference number: Z22040; MUID:97400619; PMID:9252323

A/Accession: T42047

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1846 <KIM>

A/Cross-references: UNIPROT:O16131; UNIPARC:UPI0000077B82; EMBL:AF012437; NID:92338417; C/Genetics:

A/Gene: daf-2

Query Match
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 HDIVDGNH 130
|||
Db 62 HDIVDGNH 69

RESULT 40

A54277

transcription adaptor protein p300 - human

C/Species: *Homo sapiens* (man)

C/Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004

C/Accession: A54277; S60344

R/Eckner, R.; Ewen, M.E.; Newsome, D.; Gerdes, M.; Decaprio, J.A.; Lawrence, J.B.; Livli

Genes Dev. 8, 869-884, 1994

A/Title: Molecular cloning and functional analysis of the adenovirus E1A-associated 300-

A/Reference number: A54277; MUID:95011587; PMID:7523245

A/Accession: A54277

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-2414 <ECK>

A/Cross-references: UNIPROT:Q09472; UNIPARC:UPI0000131009; GB:U01877; NID:9495300; PIDN:

A/Note: In the authors' translation 941-Ser is shown after 961 and consequently, residue

R/Lundblad, J.R.; Kwok, R.P.S.; Laurance, M.E.; Harter, M.L.; Goodman, R.H.

Nature 374, 85-88, 1995

A/Title: Adenoviral E1A-associated protein p300 as a functional homologue of the transcr

A/Reference number: S60344; MUID:95174889; PMID:7870179

A/Accession: S60344

A/Status: preliminary

A/Molecule type: protein

A/Residues: 552-660 <LUN>

A/Cross-references: UNIPARC:UPI00001782EB

C/Genetics:

A/Gene: GDB:EP300

A/Cross-references: GDB:9862958; OMIM:502700

A/Map position: 22q13.2-22q13.2

C/Superfamily: transcription coactivator CRAB-binding protein; bromodomain homology

C/Keywords: phosphoprotein; transcription; zinc finger

F/1075-1132/Domain: bromodomain homology

F/89,507,1136,1295,1497,1834,1977,2062,2220/Binding site: phosphate (Ser) (covalent) (by

F/1734/Binding site: phosphate (Ser) (covalent) (by protein kinase A) (status predicted

Query Match
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 144 QEEERRR 151
|||
Db 1522 QEEERRR 1529

RESULT 41

T42993

probable spectrin beta chain - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004

C/Accession: T42993

R/Austin, J.; Pralits, V.; McKeown, C.

submitted to the EMBL Data Library, March 1998

A/Description: Sma-1 encodes a BH-spectrin homolog required for *C. elegans* morphogenesis

A/Reference number: Z22279

A/Accession: T42993

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-4063 <AUS>

A/Cross-references: UNIPROT:O02425; UNIPARC:UPI0000110113; EMBL:AF053496; PIDN:AAC08577; C/Genetics:

A/Gene: sma1

Query Match 3.6%; Score 8; DB 2; Length 4063;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 QKXFTKX 50
 |||||
 DB 53 QKXFTKX 60

RESULT 42

hypothetical protein R31.1 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T23630; T24242

R/Kershaw, J.
 submitted to the EMBL Data Library, November 1996
 A/Reference number: Z19774

A/Accession: T23630

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-4101 <WIL>

A/Cross-references: UNIPROT:002425; UNIPARC:UPI000017CF3B; EMBL:Z81570; PIDN:CAB04608.1;
 A/Experimental source: clone K12G11

R/Lennard, N.
 submitted to the EMBL Data Library, July 1996

A/Reference number: Z19862

A/Accession: T24242

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-4101 <W12>

A/Cross-references: UNIPARC:UPI000017CF3B; EMBL:Z75956; PIDN:CAB00130.1; GSPDB:GN00023;
 A/Experimental source: clone R31

C/Genetics:

A/Gene: CBSP:R31.1

A/Map position: 5

A/Introns: 29/3; 89/3; 184/3; 312/3; 812/2; 1461/3; 1531/3; 2050/2; 2413/2; 2466/2; 3008

Query Match

Best Local Similarity 100.0%; Pred. No. 59;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 QKXFTKX 50
 |||||
 DB 75 QKXFTKX 82

RESULT 43

165646

Duchenne muscular dystrophy protein - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 09-Jul-2004

C/Accession: I65646

R/Chelly, U.; Gligenkranz, H.; Lambert, M.; Hamard, G.; Chafey, P.; Recan, D.; Katz, P.
 Cell 63, 1239-1248, 1990

A/Title: Effect of dystrophin gene deletions on mRNA levels and processing in Duchenne's
 A/Reference number: I52806; MUID:91084852; PMID:2261642

A/Accession: I65646

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-125 <RES>

A/Cross-references: UNIPROT:Q14174; UNIPARC:UPI000007156E; GB:M63075; NID:G181591; PIDN:
 C/Genetics:

A/Gene: GDB:DMD

OY 124 DIVDGNH 130
 |||||
 DB 61 DIVDGNH 67

RESULT 44

hypothetical protein YUL124c - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein J0714

C/Species: Saccharomyces cerevisiae

C/Date: 05-May-1995 #sequence_revision 08-Sep-1995 #text_change 09-Jul-2004
 C/Accession: S56905

R/Ciepluch, C.; Korde, E.; Pujol, A.; Jauniaux, J.C.
 submitted to the Protein Sequence Database, September 1995

A/Reference number: S56891

A/Accession: S56905

A/Molecule type: DNA

A/Residues: 1-172 <CZ1>

A/Cross-references: UNIPROT:P47017; UNIPARC:UPI00000530D0; EMBL:Z49399; NID:G1008319; PT
 C/Genetics:

A/Gene: SGD:LSM1; MIPS:YUL124c

A/Cross-references: SGD:S0003660

A/Map position: 10L

C/Superfamily: hypothetical protein YUL124c

Query Match 3.2%; Score 7; DB 2; Length 172;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 69 LRDRML 75
 |||||
 DB 58 LRDRML 64

RESULT 45

glucose-1-phosphate thymidyltransferase [imported] - Halobacterium sp. NRC-1

C/Species: Halobacterium sp. NRC-1

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C/Accession: B84165

R/Ng, W.V.; Kennedy, S.P.; Mahatras, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laeky, S.
 ; Leithausen, B.; Keller, K.; Cruz, R.; Danison, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebdhardt, H.; Lowe, T.M.; Li
 A/Title: Genome sequence of Halobacterium species NRC-1

A/Reference number: A84160; MUID:20504483; PMID:11016950

A/Accession: B84165

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-218 <STO>

A/Cross-references: UNIPROT:Q9HSW4; UNIPARC:UPI0000063526; GB:AE004437; NID:G10579698; ;
 C/Genetics:

A/Gene: grade

Query Match 3.2%; Score 7; DB 2; Length 218;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 140 VRRQOE 146
 |||||
 DB 92 VRRQOE 98

RESULT 46

hypothetical protein B21J21.320 [imported] - Neurospora crassa

C/Species: Neurospora crassa

C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 18-Aug-2000

C/Accession: T49719

R/Schulte, U.; Algen, V.; Hohnsels, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
 submitted to the Protein Sequence Database, May 2000

A:Reference number: Z25022
 A:Accession: F69844
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-226 <SCH>
 A:Cross-references: UNIPARC:UPI0000179D83; EMBL:AL356172; GSPDB:GN00116; NCSP:B23121.320
 A:Experimental source: BAC clone B23121; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B23121.320
 A:Intron: 90/1
 C:Superfamily: Neurospora crassa hypothetical protein B23121.320

Query Match 3.2%; Score 7; DB 2; Length 226;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 146 EERRRR 152
 |||||
 DB 38 EERRRR 44

RESULT 47

AB0957
 conserved hypothetical protein STY934 [imported] - Salmonella enterica subsp. enterica
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AB0957
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AB0957
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-251 <PAR>
 A:Cross-references: UNIPARC:UPI000005A692; GB:AL513382; PIDN:CAD03150.1; PID:G16504785;
 C:Genetics:
 A:Gene: STY934

Query Match 3.2%; Score 7; DB 2; Length 251;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MELORTS 7
 |||||
 DB 103 MELORTS 109

RESULT 48

F69844
 conserved hypothetical protein yjbn - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 31-Dec-2004
 C:Accession: F69844
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte, C.; Bron, P.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capiano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Eutian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Geller, J.; Harwood, C.R.; Henaui, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.; Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Luthers, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudaga, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portefeld, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Schuster, S.; Schuster, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Serot, A.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: F69844
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-266 <KUN>
 A:Cross-references: UNIPROT:O31612; UNIPARC:UPI000006021B; GB:Z99110; GB:AL009126; NID:5
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yjbn
 C:Superfamily: ATP-NAD(H) kinase

Query Match 3.2%; Score 7; DB 1; Length 266;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 SKGDQVS 180
 |||||
 DB 7 SKGDQVS 13

RESULT 49

G97246
 azaC-type DNA-binding domain-containing protein, transcription regulator [imported] - C]
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
 C:Accession: G97246
 R:Noelling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daily, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4823-4838, 2001
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum
 A:Reference number: A96900; MUID:21359325; PMID:21359325
 A:Accession: G97246
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-279 <KUN>
 A:Cross-references: UNIPROT:O97FC2; UNIPARC:UPI00000CA649; GB:AE001437; PIDN:AAK80762.1;
 A:Experimental source: Clostridium acetobutylicum ATCC824
 C:Genetics:
 A:Gene: CAC2818

Query Match 3.2%; Score 7; DB 2; Length 279;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 114 RVHLENM 120
 |||||
 DB 160 RVHLENM 166

RESULT 50

G86155
 hypothetical protein T14P4.13 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: G86155
 R:Tholl, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: G86155
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-290 <STO>
 A:Cross-references: UNIPROT:Q9FWX9; UNIPARC:UPI00000A60B7; GB:AE005172; NID:G9972379; PI
 C:Genetics:

A;Map position: 1

Query Match 3.2%; Score 7; DB 2; Length 290;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 144 QEEERK 150
|||
Db 219 QEEERK 225

Search completed: March 4, 2006, 07:10:34
Job time : 72 secs